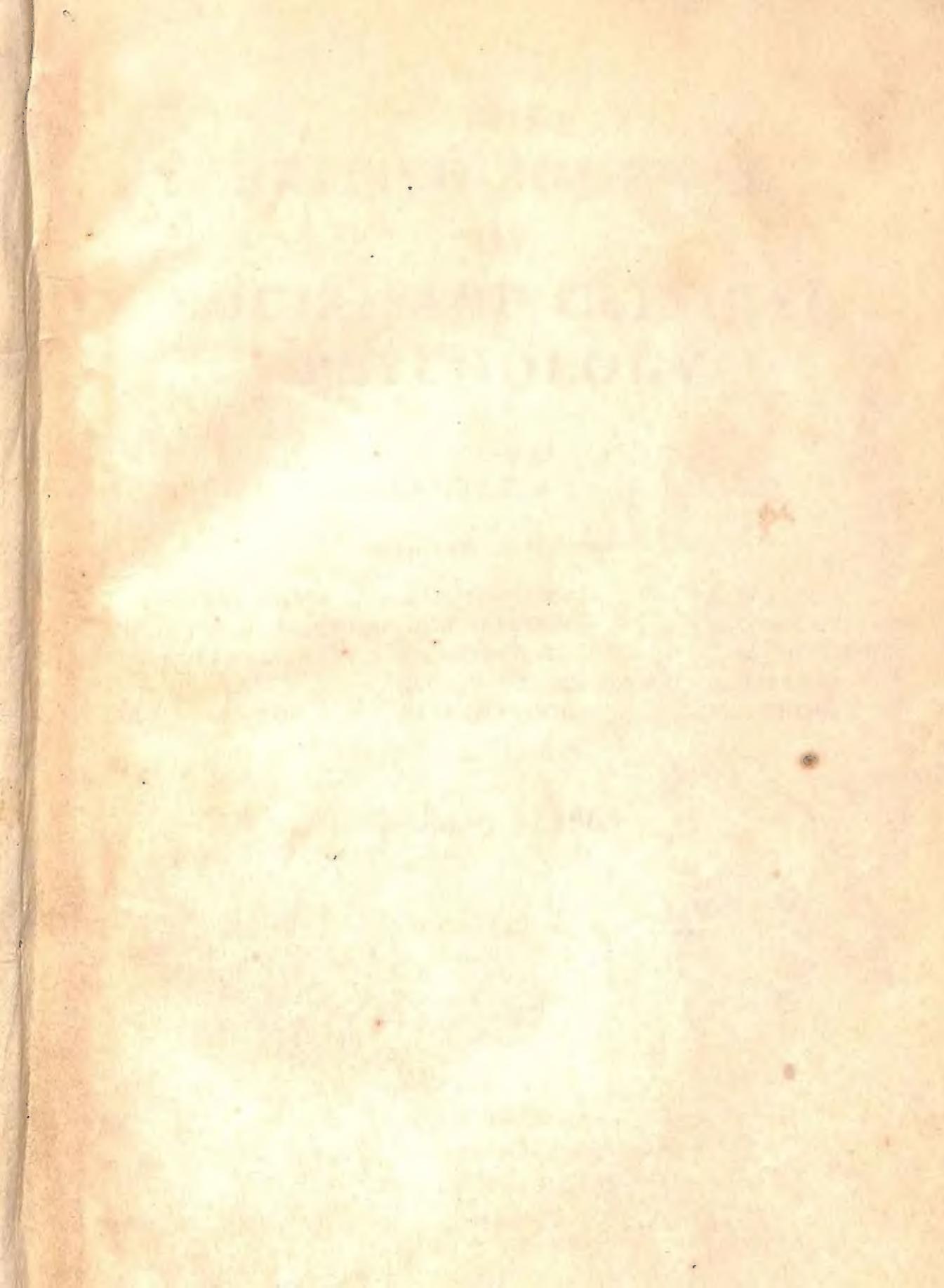




~~22/23/21~~
22.8.20







THE
BRITISH JOURNAL
OF
SOCIAL AND CLINICAL
PSYCHOLOGY

EDITED BY
MICHAEL ARGYLE & JACK TIZARD

WITH THE ASSISTANCE OF

D. RUSSELL DAVIS	A. G. HAMMER	A. R. LURIA
CECILY de MONCHAUX	D. W. HARDING	A. E. MAXWELL
H. J. EYSENCK	ALASTAIR HERON	T. M. NEWCOMB
G. A. FOULDS	HILDE T. HIMMELWEIT	R. R. SEARS
MAX HAMILTON	MARIE JAHODA	E. J. SHOBEN

Volume 2, 1963



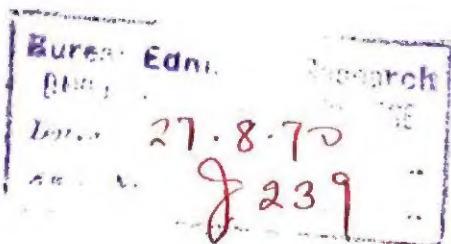
CAMBRIDGE

Published for the British Psychological Society
AT THE UNIVERSITY PRESS

1963

PUBLISHED BY

THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London N.W.1.
American Branch: 32 East 57th Street, New York 22, N.Y.
West African Office: P.O. Box 33, Ibadan, Nigeria



CONTENTS OF VOLUME 2

PART I. FEBRUARY 1963

JACKSON, DOUGLAS N. and MESSICK, SAMUEL. Individual Differences in Social Perception	<i>page</i> 1
VENESS, THELMA and BRERLEY, DOROTHY W. Forming Impressions of Personality: Two Experiments	11
POPPLETON, PAMELA K. and PILKINGTON, G. W. The Measurement of Religious Attitudes in a University Population	20
HERMELIN, BEATE and O'CONNOR, N. The Response and Self-Generated Behaviour of Severely Disturbed Children and Severely Subnormal Controls	37
ABI RAFI, A. A Note on the Cultural Aspects of the WAIS Vocabulary Subtest in Relation to British Mental Patients	44
EYSENCK, S. B. G. and EYSENCK, H. J. On the Dual Nature of Extraversion	46
ROBINSON, J. O. A Study of Neuroticism and Casual Arterial Blood Pressure	56
INGLIS, JAMES and MACKAY, H. A. A Method for Testing 'Simulated' Visual Field Defects	65
RICHARDSON, ALAN. A Reply to Dr. Emery	67
Book Reviews	71

PART 2. JUNE 1963

SHAPIRO, M. B., MARKS, I. M. and FOX, B. A Therapeutic Experiment on Phobic and Affective Symptoms in an Individual Psychiatric Patient	<i>page</i> 81
VENABLES, P. H. Changes Due to Noise in the Threshold of Fusion of Paired Light Flashes in Schizophrenics and Normals	94
LOVIBOND, S. H. Conceptual Thinking, Personality and Conditioning	100
SARBIN, THEODORE R. and ANDERSEN, MILTON L. Base-Rate Expectancies and Perceptual Alterations in Hypnosis	112
OLDFIELD, R. C. Individual Vocabulary and Semantic Currency: A Preliminary Study	122
KNOWLES, J. B. Acquiescence Response Set and the Questionnaire Measurement of Personality	131
VENABLES, ETHEL C., Social Differences among Day-release Students in Relation to Their Recruitment and Examination Success	138
Book Reviews	153

PART 3. OCTOBER 1963

RUTTER, MICHAEL, KORN, SAM and BIRCH, HERBERT G. Genetic and Environmental Factors in the Development of Primary Reaction Patterns'	page 161
WOODWARD, MARY. Early Experience and Behaviour Disorders in Severely Subnormal Children	174
HAMMER, A. G. Neuroticism, Extraversion and Hypnotizability	185
CENTERS, LOUISE and CENTERS, RICHARD. Perceptual Sensitivity and Defense in Amputee Children	192
SIEGMAN, ARON WOLFE. A Cross-cultural Investigation of the Relationship Between Introversion-Extraversion, Social Attitudes and Anti-social Behaviour	196
MUSGROVE, F. Inter-generation Attitudes	209
CATTELL, RAYMOND B. The Meaning and Measurement of Neuroticism and Anxiety: Supplement to a Review	224
Book Reviews	226

INDEX TO AUTHORS

ABI RAFI, A.	44	MARKS, I. M.	81
ANDERSEN, MILTON L.	112	MESSICK, SAMUEL	1
BIRCH, HERBERT G.	161	MUSGROVE, F.	209
BRIERLEY, DOROTHY W.	11	O'CONNOR, N.	37
CATTELL, RAYMOND B.	224	OLDFIELD, R. C.	122
CENTERS, LOUISE	192	PILKINGTON, G. W.	20
CENTERS, RICHARD	192	POPLETON, PAMELA K.	20
EYSENCK, H. J.	46	RICHARDSON, ALAN	67
EYSENCK, S. B. G.	46	ROBINSON, J. O.	56
FOX, B.	81	RUTTER, MICHAEL	161
HAMMER, A. G.	185	SARBIN, THEODORE R.	112
HERMELIN, BEATE	37	SHAPIRO, M. B.	81
INGLIS, JAMES	65	SIEGMAN, ARON WOLFE	196
JACKSON, DOUGLAS N.	1	VENABLES, ETHEL C.	138
KNOWLES, J. B.	131	VENABLES, P. H.	94
KORN, SAM	161	VENESS, THELMA	11
LOVIBOND, S. H.	100	WOODWARD, MARY	174
MACKAY, H. A.	65		

INDEX OF PUBLICATIONS REVIEWED

ALLEN, CLIFFORD. <i>A Textbook of Psychosexual Disorders</i> 75	HENDERSON, SIR DAVID & BATELOR, IVOR R. C. <i>Henderson and Gillespie—Textbook of Psychiatry</i> (9th edn. revised) 238
ALLPORT, GORDON W. <i>Pattern and Growth in Personality</i> 153	HENLE, MARY (ed.). <i>Documents of Gestalt Psychology</i> 157
ANDERSON, NELS. <i>Work and Leisure</i> 77	HOMANS, G. C. <i>Social Behaviour. Its Elementary Forms</i> 227
ANDRY, R. G. <i>The Short-Term Prisoner</i> 230	HUMPHREY, GEORGE & ARCYLE, MICHAEL (eds.). <i>Social Psychology Through Experiment</i> 155
ARMSTRONG, D. M. <i>Bodily Sensations</i> 237	KASTEL, JEAN. <i>Casework in Child Care</i> 240
BIRNEY, R. C. & TEEVAN, R. C. (eds.). <i>Measuring Human Motivation</i> 159	McNEMAR, QUINN. <i>Psychological Statistics</i> 79
BLISS, E. L. (ed.). <i>Roots of Behaviour</i> 231	MASLOW, ABRAHAM H. <i>Toward a Theory of Being</i> 236
BLUM, GERALD S. <i>A Model of the Mind</i> 234	NEWCOMB, THEODORE M. <i>The Acquaintance Process</i> 229
BULLOUGH, GEOFFREY. <i>Mirror of Minds</i> 78	O'CONNOR, NEIL & HERMELIN, BEATE. <i>Speech and Thought in Severe Subnormality</i> 232
CANDLAND, K. (ed.). <i>Emotion: Bodily Change</i> 240	RAY, ISAAC. <i>Overholser, WINFRED</i> (ed.). <i>A Treatise on the Medical Jurisprudence of Insanity</i> 234
CATTELL, R. B. & SCHEIER, I. H. <i>The Measurement of Neuroticism and Anxiety</i> 73	RICHTER, D., TANNER, J. M., TAYLOR, LORD & ZANGWILL, O. L. (eds.). <i>Aspects of Psychiatric Research</i> 154
Farber, Seymour M. & Wilson, Roger H. L. (eds.). <i>The Control of the Mind</i> 78	SHERIF, MUZAFER & HOVLAND, CARL I. <i>Social Judgment. Assimilation and Contrast Effects in Communication and Attitude Change</i> 228
FRASER, J. MUNRO. <i>Industrial Psychology</i> 160	SMELESER, NEIL J. <i>Theory of Collective Behaviour</i> 232
GELDARD, F. A. <i>Fundamentals of Psychology</i> 156	
GEORGE, F. H. <i>Cognition</i> 157	
GETZELS, JACOB W. & JACKSON, PHILIP W. <i>Creativity and Intelligence: Explorations with Gifted Students</i> 71	
GLUECK, SHELDON & GLUECK, ELEONOR. <i>Family Environment and Delinquency</i> 159	
GORDON, IRA J. <i>Human Development: From Birth through Adolescence</i> 236	
HAIDER, DR. MANFRED. <i>Ermüdung, Beanspruchung und Leistung</i> 75	

Index of Publications Reviewed

SROLE, LEO, LANGNER, THOMAS S., MICHAEL, STANLEY, T., OPLER, MARVIN K. & RENNIE, THOMAS A. C. *Mental Health in the Metropolis: the Midtown Manhattan Study* (vol. 1) 72

STARK, WERNER. *The Fundamental Forms of Social Thought* 238

TAFT, RONALD & DOCZY, A. G. *The Assimilation of Intellectual Refugees in Western Australia: with special reference to Hungarians* 239

TRAPP, PHILIP E. & HIMELSTEIN, PHILIP (eds). *Readings on the Exceptional Child: Research and Theory* 235

VALENTINE, C. W. *The Experimental Psychology of Beauty* 227

VENESS, THELMA. *School Leavers—Their Aspirations and Expectations* 158

WELFORD, A. T., ARGYLE, MICHAEL, GLASS, D. V. & MORRIS J. N. (eds.). *Society: Problems and Methods of Study* 76

Individual Differences in Social Perception*

By DOUGLAS N. JACKSON
Pennsylvania State University

AND SAMUEL MESSICK
Educational Testing Service

This review highlights the importance of studying the role of the perceiver in social perception. While noting that there is often a strong temptation to simplify such cognized social attributes as status, dominance and leadership into single orderings, it is emphasized that measurement procedures must faithfully represent the complexity inherent in the domain if precision and understanding are to be achieved. Alternative multidimensional scaling models, including procedures for taking account of differently-structured points of view, are described in terms of their relevance to social perception.

The aim of the present paper is twofold: (a) to call attention to the important role that the perceiver plays in determining the manner in which other people and social events are organized, and (b) to describe possible linkages between mathematical scaling models on the one hand and concrete problems in the areas of person cognition and social perception on the other.

In reviewing the extensive literature in social perception accumulated over the past fifteen years, we were impressed with both its scope and its vigour. We are convinced, as Bruner & Tagiuri (1954) have noted, that an understanding of the processes by which social events are cognized should provide an essential foundation for theory construction in interpersonal relations and social psychology. To be sure, the issues are very complex, and no single method is likely to solve all of the vexing problems confronting the student of person cognition. We do believe, however, that certain recently developed psychometric procedures are available which may well provide a basis for a methodologically and psychologically sound analysis of certain complex processes operating in social perception.

The very complexity inherent in the domain of person cognition has provided one of the most fertile grounds for incisive critiques of the methodology used by investigators attempting to proceed with the solution of important theoretical problems (cf. Abelson, 1960; Campbell, 1955; Campbell & Fiske, 1959; Cline & Richards, 1960; Cronbach, 1955, 1958; Cronbach & Gleser, 1953; Crow & Hammond, 1957; Gage & Cronbach, 1955; Gulliksen, 1958; La Forge, 1959; Luchins, 1948; Lykken, 1956; Quereshi, 1959; Taft, 1955; Thorndike & Stein, 1937; Wishner, 1960). While these able critics have provided us with a more mature outlook and an improved methodology, the rogues gallery of artifacts and errors uncovered by these psychological sleuths is cause for even the most intrepid and optimistic investigator to tread cautiously, if not to become immobilized. The

* Prepared for a symposium on *The Perceiver in Social Perception*, American Psychological Association, New York, 1961.

investigator wishing to study social perception must cope not only with theoretical and experimental challenges unique to his problem, but must also take into account, among other pitfalls, possible method variance, confounded or experimentally dependent scores, metric assumptions required in data analysis, response sets, rating biases, oversimplified general distance and similarity indices, interactions between judge and concept or person, nonlinearity, asymmetry, implicit assumptions about what is relevant to the judge or the domain, and the like.

While we value highly the advice and criticism of able ambassadors from the stern land of *Psychometrika* (to borrow Cronbach's (1954) analogy), as well as from the verdant countryside of social psychology—and have even on occasion attempted to travel this road ourselves—we are not now attempting to serve as nomadic critics seeking to demonstrate that some previous research can be accounted for by this or by that artifact. Rather, our goal is to make a few suggestions as to how the perceiver of social events may be studied in a manner which attempts to do justice to his complex and sometimes disconcertingly idiosyncratic cognitive organization.

Our central thesis is that for an accurate and full understanding of person perception, measurement procedures must bear some realistic relation to the complexity of the domain; the apparently infinite varieties of perceived differences among people can indeed be structured and ordered, but they cannot with impunity be forced into a mold of preconceived simplicity.

We do not deny that there is strong temptation to simplify. De Soto (1961) has argued persuasively, with evidence, that the predilection for single orderings, for a unidimensional schemata about other people, is embedded not only in colloquial 'lay personality theory', but is woven deeply even into the thinking of the social scientist in his quest for simple orderings of social class, leadership, dominance and the like. The social scientist seeks perhaps to avoid cognitive strain as much as other people, but may be more dismayed than the average person to find that indices of social class—power, wealth, prestige, education—are imperfectly correlated; that leadership is a function of no single attribute and is to a degree situationally bound; and that dominance relationships are sometimes intransitive. In the area of person perception one might legitimately inquire of the investigator to what extent an appeal to the law of parsimony is a defensive attempt to alleviate cognitive dissonance.

We should like to propose one solution to the dual problem of rigorously representing ordered schemata and realistically allowing for multiple orderings both in the characteristics of the people being judged and in the points of views of the judges about these characteristics. Before discussing the model—an extension of multidimensional scaling—and its applications to data in person perception and trait implication, it is desirable to show its ancestry and derivation.

Table 1 is a fourfold representation of certain scaling models suggested to us by Gulliksen. In the upper left-hand quadrant are the traditional unidimensional scaling methods—paired comparisons, successive intervals and the like (cf. Torgerson, 1958)—yielding simple orderings, or perhaps interval scales representing a kind of average of the ratings of all judges along a single dimension. The

methods assume a single point of view and a unitary stimulus attribute. If an investigator wished to order a group of people in terms of their perceived status, for example, this method would yield a single dimension, an ordering with Duncan, Dunlap, Dunstone and McDuff represented as points upon the scale. But one might question this imposed simplicity upon the concept of status. It might be argued that while Duncan is of noble birth, he recently had to auction his castle to pay the grocer; McDuff is wealthy, but he earns his living making book, and so on. What is then required is a method to appraise the degree to which the imposed unidimensionality assumed for the stimulus attribute accurately reflects the perceived stimulus differences.

Table 1. *Classification of certain unidimensional and multidimensional scaling models*

Paired comparisons Successive intervals Equal-appearing intervals	Tucker's vector model for paired comparisons Factor analysis of trait ratings
Multidimensional successive intervals Multidimensional method of triads	Tucker-Messick individual differences model of multidimensional scaling

The lower left-hand quadrant of Table 1 includes methods for evaluating such an assumption of a single ordering for the stimulus attribute. These methods—traditional multidimensional scaling procedures (Messick, 1956b; Torgerson, 1958)—yield an average of the perceived distances among a set of stimuli, which in the present context might include other people or personality traits. The procedures entail the obtaining of judgments of similarity or difference between pairs of stimuli, constructing a matrix of relative distances between all possible stimulus pairs, meeting assumptions of ratio measurement by transforming the relative distances to absolute distances by solving for an additive constant, and then factoring a matrix of scalar products (Messick & Abelson, 1956; Torgerson, 1958). The distances between stimuli are thus represented in a Euclidean space, the dimensionality of which corresponds to the number of different ways the stimuli are perceived to differ. The points in the space represent the stimuli, and their projections on the obtained dimensions are scale values for the various attributes. The mutual distances among the points correspond to the psychological distances between the stimuli, be they psychophysical attributes (Torgerson, 1951; Messick, 1956c; Shepard, 1958; Indow & Karazawa, 1960); other people (Jackson, Messick & Solley, 1957; Morton, 1959); facial expressions (Abelson & Sermat, in press); or attitude statements (Abelson, 1954-55; Messick, 1956a). One of us (Jackson, 1958) has proposed that a rigorous analysis of Asch's (1946, 1952) formulation on trait implication might well be attempted with multidimensional scaling because of the apparently close relationship between trait inference and the assumptions regarding psychological distance in the model. Traits appearing close together in a multidimensional space would be expected to imply or be similar to one another; a 'central' trait in terms of the model would be one with substantial loadings on

one or more large dimensions, etc. This formulation is now being evaluated empirically by Starr at Princeton and by Walters at Pennsylvania State University. It should provide a more precise delineation of the theoretical alternatives proposed by Asch, and might complement Wishner's (1960) recent correlational approach to the study of personality impression formation.

It should be noted that this model of multidimensional scaling assumes that every judge perceives the stimuli in the same way—that while the stimuli are allowed to vary in many ways, there is only one point of view about stimulus similarity represented in the obtained average distances. If one subset of judges was cognitively complex in Bieri's (1955) terms, seeing people differing in many ways, and another subset of judges was cognitively simple, then the obtained space, representing an average, might provide more dimensions than would characterize certain of the judges and possibly distort the relationships in general. Where judges can be assumed to possess equivalent cognitive structures, or where interest is primarily in mapping the stimulus domain, the model should prove quite satisfactory as long as scaling assumptions are met. But where interest is centered not only upon the stimuli but upon the unique processes of particular classes of judges, then the representation will be accurate for all judges only to the degree that judges are homogeneous. Such judge homogeneity is unlikely in person-perception.

The problem of representing different points of view or schools of thought among judges takes us to a different multidimensional scaling approach, represented in the upper right-hand quadrant of Table 1. This method, a vector model for paired comparisons developed by Tucker (1955, 1960), assumes that each judge rates the stimuli with respect to a one-dimensional attribute, such as preference, but that the judges may differ among themselves in their perceived orderings—multidimensionality is thus sought among the judges rather than the stimuli. Thus, if we had a set of stimulus persons varying in the components of status—wealth, prestige, education, power, skill, etc.—judges, in ordering the perceived status of these persons, might differentially weight the components; their resulting different orderings would permit us to isolate their respective points of view in regard to status. A point of view is represented by a vector extending in some particular direction within the multidimensional space of stimulus objects. Stimulus projections on each vector represent scale values for each point of view. A similar rationale underlies the factor analysis of ratings of traits or people with respect to a single, but possibly complex or global criterion, such as status or desirability (Messick, 1960). These vector models thus isolate dimensions representing consistent individual viewpoints from ratings of single stimuli with respect to some specified attribute; whereas the preceding distance models isolate dimensions of perceived variation from ratings of pairs of stimuli with respect to similarity.

But the question arose as to whether it is entirely accurate, on the one hand, to average over all judges to obtain a multidimensional space of stimulus attributes or, on the other hand, to seek diversity among judgmental viewpoints while assuming one-dimensional perceptions for each judge. Tucker & Messick (1960) solved this problem by combining the two approaches, as indicated in the lower

right-hand quadrant of Table 1. This new approach takes account of the stimulus complexity inherent in the cognized social objects, while also permitting the isolation of consistent individual viewpoints in structuring the perceived stimulus properties. Thus, a separate multidimensional space is derived for *each* different viewpoint about stimulus similarity that is uncovered, and relationships among the viewpoints and with any measured characteristics of the judges are described in a multifactor space of perceiver differences.

The nature of this model will be illustrated by the application of the method to two problems in social perception: a study of the perception of prominent political figures (Messick, 1961; Tucker & Messick, 1960) and an experimental study of the simulation of international relations (Driver, 1961).

In the politics study (Tucker & Messick, 1960), judges—approximately equally divided among the four categories of (a) liberal Democrat, (b) conservative Democrat, (c) liberal Republican, and (d) conservative Republican—rated the similarity between all pairs of political figures, as listed in Table 2. Analysis

Table 2. *Political relations in three perceptual spaces*
(Tucker & Messick, 1960)

	Idealized individual		
	A	B	C
	I	I	II
Thomas Dewey	—	+	—
Senator Dirksen	—	+	—
Dwight Eisenhower	—	+	—
Senator McCarthy	—	+	++
Richard Nixon	—	+	—
Senator Taft	—	+	—
Senator Douglas	—	—	—
Senator George	—	—	—
Senator Kefauver	—	—	—
Franklin D. Roosevelt	—	—	—
Adlai Stevenson	—	—	—
Governor Talmadge	—	—	—
Harry Truman	—	—	—
General MacArthur	—	+	—
Alger Hiss	+++	(—)	++
Henry Wallace	—	(—)	++
Chiang Kai-Shek	+	○	—
Adolf Hitler	+++	○	++
Jawaharlal Nehru	+	○	—
Joseph Stalin	+++	○	++

In this space, there are, possibly, seven significant dimensions, only one of which approaches being as large as the major dimensions in the preceding spaces for idealized individuals A and B. The structure in these seven dimensions appears rather complex.

revealed three dimensions of individual differences, representing three separate multidimensional points of view, which after rotation appeared to relate moderately to political ideology. These points of view may be represented by idealized individuals, as indicated by the points A, B, and C in Figure 1. Idealized individual A revealed a markedly one-dimensional space, with a strong evaluative component (Table 2). Idealized individual B's space had two large dimensions, apparently reflecting respectively a Republican-Democratic distinction and an evaluative

attribute. Idealized individual C's space was much more complex, revealing many subtle distinctions among the political figures, quite different from the simple differences noted for the first two types. This study raises many intriguing questions as to the determinants of perceiving political figures in a complex or a simple way, the personality correlates of certain dimensions, stability of the obtained

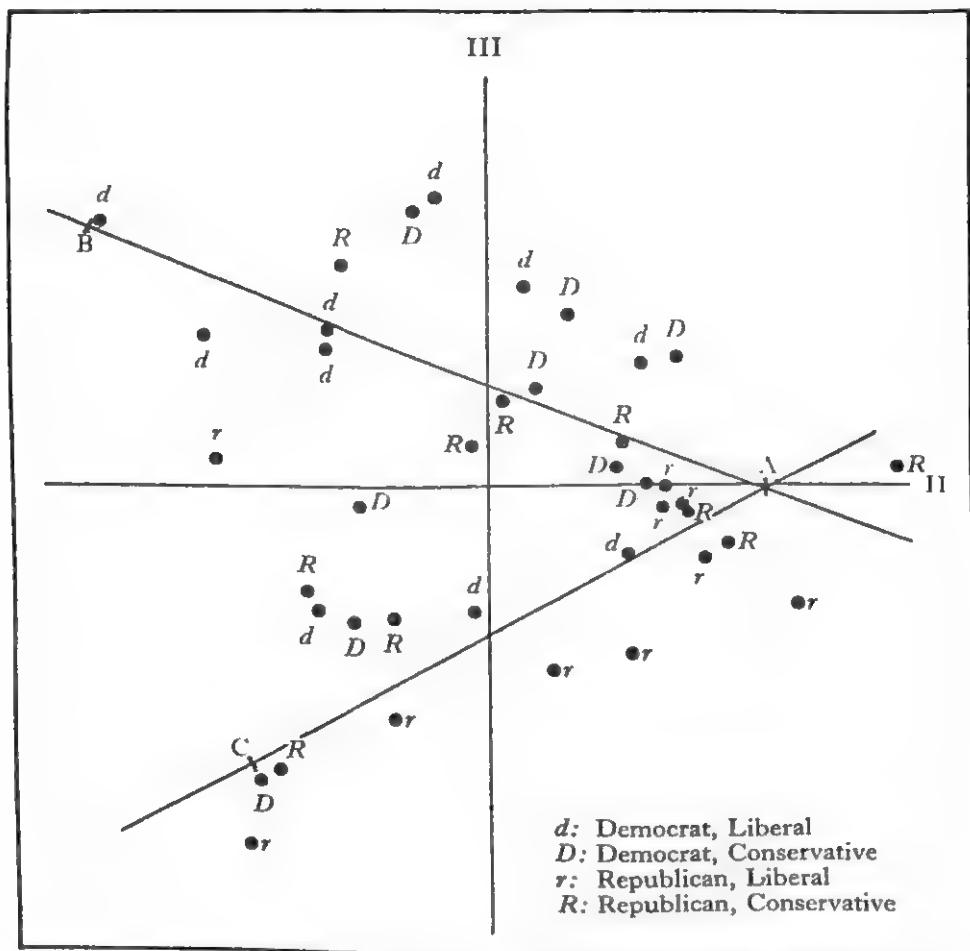


Fig. 1. Factor Structure of Individuals. Politics Interpoint Distance Data*

* All individuals had approximately equal large positive coefficients on the principal factor I. The graph gives the plot between coefficients on principal factors II and III (Tucker & Messick, 1960).

spaces, and the effect of certain experiences and information upon the perceptual structure. The present scaling model should permit a more refined evaluation of such questions than previously was possible.

Driver (1961) has explored the possibility of combining the individual differences model of multidimensional scaling with experimental manipulations in a simulated international relations study of 'nations' at peace and at war (Guetzkow, 1959). The findings were quite dramatic in certain particulars: For example, the chief executive officer of a country appeared to maintain an articulated, integrated,

'open', multidimensional view of other 'nations' to the degree that his personality questionnaire scores were low in concreteness and authoritarianism or acquiescence. While it might perhaps be premature to call together the Secretaries of State of the

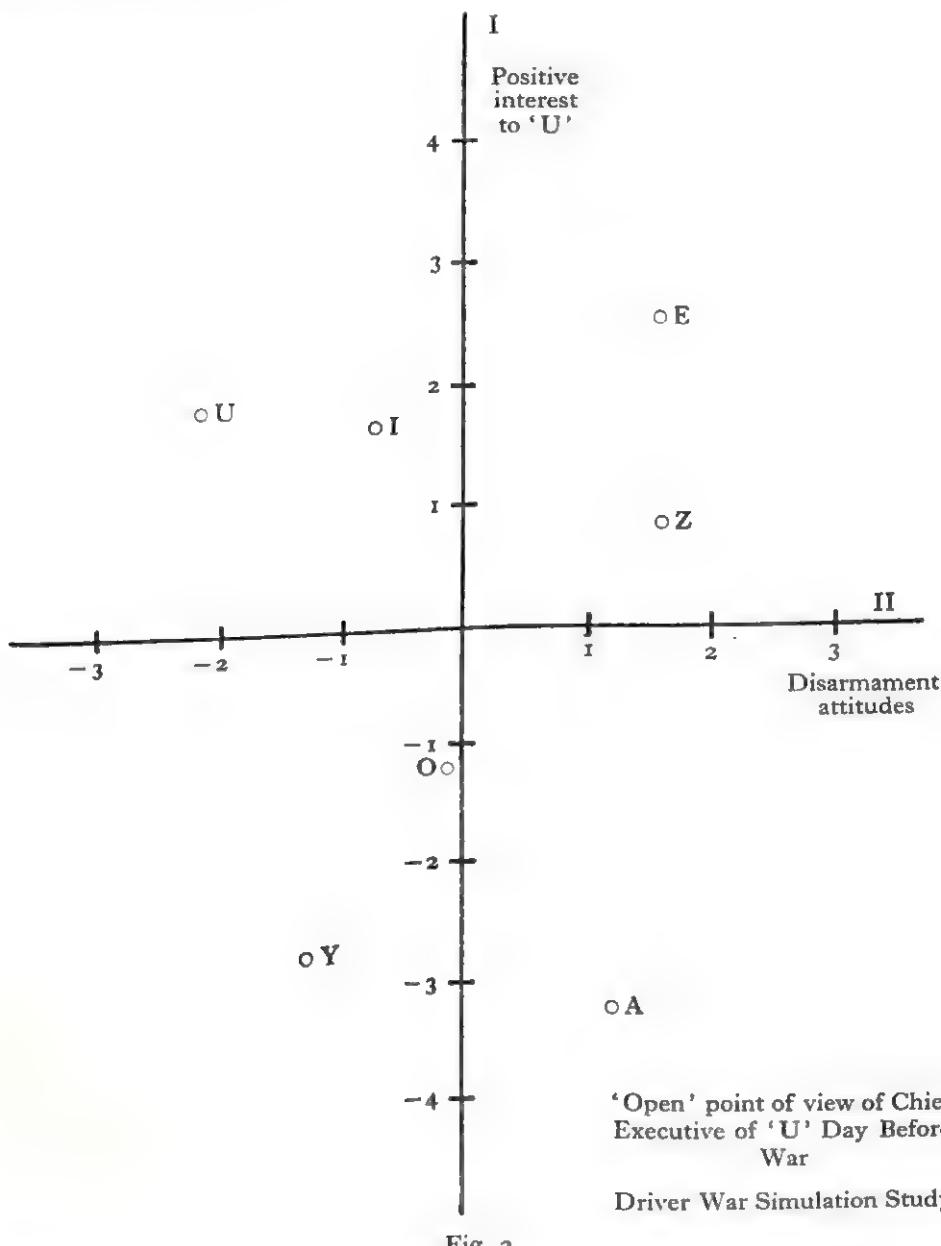


Fig. 2

major world nations for personality assessment, these results are nevertheless suggestive.

Another important finding of the Driver study was the constricting effect of the stress generated by an open outbreak of hostilities upon the multidimensional perceptions of statesmen. Subjects maintained a differentiated view of other countries during peace (Figure 2), but the declaration of war resulted in marked

simplification (Figure 3). Previously complex multidimensional spaces assimilated to a dominant evaluative or affiliative dimension. These results suggest that the presentation of the scaling technique at various intervals, combined with inter-

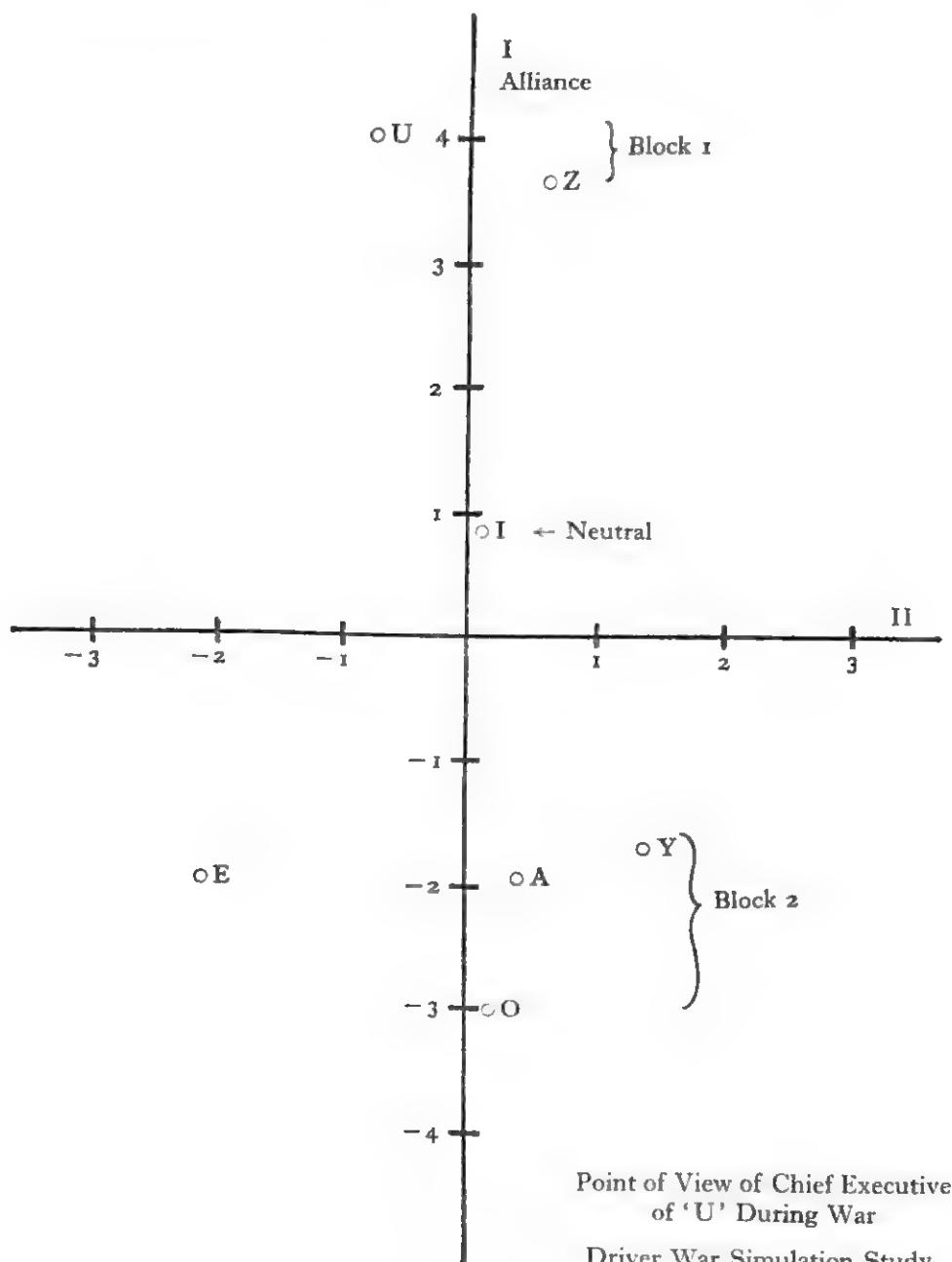


Fig. 3

vening experimentally imposed changes in the social situation, is a promising strategy for studying the dynamics of social perception.

In much research in person cognition, characteristics and processes of the perceiver are relegated to 'error' variance. In this review we have emphasized the

importance of studying the role of the perceiver, as have many previous investigators (Benedetti & Hill, 1959; Crockett, 1961; Cronbach, 1958; Jackson, 1958; Mayo, 1959; Murphy, 1947; Taft, 1955). As we have attempted to demonstrate, the emergence of different viewpoints in the studies reviewed strongly suggests that investigators of social perception should generalize to consistent individual differences whenever they appear in the perception of social objects, rather than to a generalized 'average person' who may not, in fact, represent anyone.

REFERENCES

ABELSON, R. P. (1954-55). A technique and a model for multidimensional attitude scaling. *Publ. Opin. Quart.* **18**, 405-418.

ABELSON, R. P. (1960). Scales derived by consideration of variance components in multi-way tables. In H. Gulliksen & S. Messick (Eds.) *Psychological Scaling: Theory and Applications*. New York: John Wiley & Sons, Inc. Pp. 169-186.

ABELSON, R. P. & SERMAT, V. Multidimensional scaling of facial expressions. *J. exp. Psychol.* (in press).

ASCH, S. E. (1946). Forming impressions of personality. *J. abnorm. soc. Psychol.* **41**, 258-290.

ASCH, S. E. (1952). *Social Psychology*. New York: Prentice-Hall.

BENEDETTI, D. T. & HILL, J. G. (1960). A determiner of the centrality of a trait in forming impressions of others. *J. abnorm. soc. Psychol.* **60**, 278-280.

BIERI, J. (1955). Cognitive complexity-simplicity and predictive behavior. *J. abnorm. soc. Psychol.* **51**, 263-268.

BRUNER, J. S. & TAGIURI, R. (1954). The perception of people. In G. Lindzey (Ed.) *Handbook of Social Psychology*. Vol. I. Cambridge, Mass.: Addison-Wesley. Pp. 634-654.

CAMPBELL, D. T. (1955). An error in some demonstrations of the superior social perceptiveness of leaders. *J. abnorm. soc. Psychol.* **51**, 694-695.

CAMPBELL, D. T. & FISKE, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* **56**, 81-105.

CLINE, V. B. & RICHARDS, M. (1960). Accuracy of interpersonal perception—a general trait? *J. abnorm. soc. Psychol.* **60**, 1-7.

CROCKETT, W. H. (1961). Some cognitive properties of the perceiver which affect interpersonal perception. Paper read at Amer. Psychol. Assoc., New York.

CRONBACH, L. J. (1954). Report on a psychometric mission to Clinicia. *Psychometrika*, **19**, 263-270.

CRONBACH, L. J. (1955). Processes affecting scores on 'understanding of others' and 'assumed similarity'. *Psychol. Bull.* **52**, 177-194.

CRONBACH, L. J. (1958). Proposals leading to analytic treatment of social perception scores. In R. Tagiuri & L. Petrullo (Eds.) *Person Perception and Interpersonal Behavior*. Stanford: Univer. Press. Pp. 352-379.

CRONBACH, L. J. & GLESER, C. (1953). Assessing similarity between profiles. *Psychol. Bull.* **50**, 456-474.

CROW, W. J. & HAMMOND, K. R. (1957). The generality of accuracy and response sets in interpersonal perception. *J. abnorm. soc. Psychol.* **54**, 384-390.

DE SOTO, C. (1961). The predilection for single orderings. *J. abnorm. soc. Psychol.* **62**, 16-23.

DRIVER, M. (1961). Unpublished doctoral dissertation, Princeton Univer., in preparation.

GAGE, N. L. & CRONBACH, L. J. (1955). Conceptual and methodological problems in interpersonal perception. *Psychol. Rev.* **62**, 411-422.

GUETZKOW, H. (1959). A use of simulation in inter-nation relations. *Behav. Sci.* **4**, 183-191.

GULLIKSEN, H. (1958). How to make meaning more meaningful. *Contemp. Psychol.* **3**, 115-119.

INDOW, T. & KARAZAWA, K. (1960). Multidimensional mapping of Munsell colors varying in hue, chroma, and value. *J. exp. Psychol.* **59**, 330-336.

JACKSON, D. N. (1958). The measurement of perceived personality trait relationships. Air Force Office of Scientific Research, Technical Report. Reprinted in N. Washburn (Ed.) *Values, Decisions, and Groups*. New York: Pergamon Press (in press).

JACKSON, D. N., MESSICK, S. & SOLLEY, C. M. (1957). A multidimensional scaling approach to the perception of personality. *J. Psychol.* **44**, 311-318.

LA FORGE, R. (1959). The 'generality' of the D^2 -statistic. Paper read at the Midwest Psychol. Assoc., Chicago, May, 1959.

LUCHINS, A. S. (1948). Forming impressions of personality: a critique. *J. abnorm. soc. Psychol.* **43**, 318-325.

LYKKEN, D. T. (1956). A method of actuarial pattern analysis. *Psychol. Bull.* **53**, 102-107.

MAYO, CLARA (1959). The effect of cognitive complexity on conflict resolution in impression formation. Paper read at Eastern Psychol. Assoc., Atlantic City.

MESSICK, S. (1956a). The perception of social attitudes. *J. abnorm. soc. Psychol.* **52**, 57-66.

MESSICK, S. (1956b). Some recent theoretical developments in multidimensional scaling. *Educ. psychol. Measmt.* **16**, 82-100.

MESSICK, S. (1956c). An empirical evaluation of multidimensional successive intervals. *Psychometrika*, **21**, 367-375.

MESSICK, S. (1960). Dimensions of social desirability. *J. consult. Psychol.* **24**, 279-287.

MESSICK, S. (1961). The perceived structure of political relationships. *Sociometry*, **24**, 270-278.

MESSICK, S. & ABELSON, R. P. (1956). The additive constant problem in multidimensional scaling. *Psychometrika*, **21**, 1-15.

MORTON, A. S. (1959). Similarity as a determinant of friendship: a multi-dimensional study. Princeton, N.J.: ONR Technical Report, Educational Testing Service.

MURPHY, G. (1947). *Personality*. New York: Harpers.

QUERESHI, M. Y. (1959). The generality of social perception scores. *J. abnorm. soc. Psychol.* **58**, 317-322.

SHEPARD, R. N. (1958). Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *J. exp. Psychol.* **55**, 509-523.

TAFT, R. (1955). The ability to judge people. *Psychol. Bull.* **52**, 1-23.

THORNDIKE, R. L. & STEIN, S. (1937). An evaluation of the attempts to measure social intelligence. *Psychol. Bull.* **34**, 275-285.

TORGERSON, W. S. (1951). A theoretical and empirical investigation of multidimensional scaling. Unpublished doctoral dissertation, Princeton Univer.

TORGERSON, W. S. (1958). *Theory and Methods of Scaling*. New York: John Wiley & Co.

TUCKER, L. R. (1955). Description of paired comparison preference judgments by a multidimensional vector model. Research Memorandum 55-7. Princeton, N.J.: Educational Testing Service.

TUCKER, L. R. (1960). Intra-individual and inter-individual multidimensionality. In H. Gulliksen & S. Messick (Eds.) *Psychological Scaling: Theory and Applications*. New York: John Wiley & Sons, Inc. Pp. 155-167.

TUCKER, L. R. & MESSICK, S. (1960). Individual differences in multidimensional scaling. Princeton, N.J.: Educational Testing Service Research Memorandum.

WISHNER, J. (1960). Reanalysis of 'impressions of personality'. *Psychol. Rev.* **67**, 96-112.

Manuscript received 23 December 1961

Forming Impressions of Personality: Two Experiments

BY THELMA VENESS AND DOROTHY W. BRIERLEY

Birkbeck College, University of London

Asch's experiment on forming impressions of personality on the basis of lists of traits, one trait being different for the two experimental groups, is reduplicated with certain variations. In the first experiment, the traits are embodied in a descriptive passage reproduced on tape, the variable trait being represented by changes in tone of voice. The second experiment makes only a slight modification on Asch's original experiment, and confirms his main findings: however, the data are, in part, differently analysed. The results of the two experiments are discussed in relation to theoretical points raised by Asch, some of which have been recently disputed by Wishner.

INTRODUCTION

Asch's experiments on forming impressions of personality from lists of trait names, using the technique of varying one trait only (Asch, 1946), have been followed by a number of studies using the same technique (Mensch & Wishner, 1947; Kelley, 1950) and others using different experimental situations but also investigating the way in which personality impressions are formed on the basis of similar stimulus material (Cofer & Dunne, 1952; Shapiro & Tagiuri, 1958). Subjects seem to have little difficulty in forming intelligible impressions of a hypothetical person described by a few characteristics only, suggesting that the 'implicit personality theory' (Bruner & Tagiuri, 1954) we apply when we form initial impressions of real people may owe much to systematized expectations based on relatively few cues. Even when apparently incompatible information is provided (Asch, 1946; Haire & Grunes, 1950; Gollin, 1958), many people are able to find means of reconciliation that permit them to form coherent personality impressions.

While Asch's experimental technique has largely been accepted without criticism, with one exception, certain of his theoretical points (Asch, 1946, 1952) have aroused controversy. Asch, with his preference for field rather than association explanations, spoke of the traits themselves 'seeking organization', rather than of past experience leading to expectations of certain relationships between traits (Bruner, Shapiro & Tagiuri, 1958) or of the logical relatedness of traits (Hays, 1958). Further, the principles of organization were adumbrated: the given traits organize themselves in a hierachical way, certain traits occupying a central position. The warm-cold variable was the one which best illustrated this situation, and several later studies, with one exception (Kjeldergaard & Jenkins, 1958), have also claimed that this variable takes up a central position. Until challenged by Wishner (1960), the evidence for the centrality of this variable seemed overwhelming. Hays (1958), in a study in which 'centrality' is defined, in part, in terms of implication, demonstrated that the terms 'warm' and 'cold' were ones which implied many other traits while not being implied by them. (Indeed, Shapiro & Tagiuri (1958), using a technique in which variations were made in the response list instead of the

stimulus list, found that 'warm' was seldom inferred from the given characteristics ('intelligent' and 'independent'.) Asch himself extended his analysis of the effects of 'warm' and 'cold' by studying their metaphorical uses in various languages, concluding that common functional relations are involved in metaphor, 'warm' always denoting physically and psychologically a property that brings us closer, 'cold' one that excludes and isolates (Asch, 1958). On this basis, one would expect the warm-cold variable to have a certain prepotency when we are forming first impressions, so that Wishner's refutation of its centrality in experimental impression-formation is unexpected and worthy, therefore, of close scrutiny.

The experiments to be reported were conducted before they could benefit from Wishner's work. The findings will be evaluated, however, in the light of his discussion.

EXPERIMENT 1

The one criticism made of Asch's procedure, by Luchins (1948), was that it was divorced from reality in dealing only with trait names and hypothetical persons. Kelley (1950) used the Asch technique but introduced real persons. The present experiment attempts to meet Luchins' criticism by presenting stimulus material in the form of tape-recordings of a human voice, whereas the warm-cold variation is represented in the manner of speaking.

Subjects

Sixty women students, in the second year of a course of training to be non-graduate teachers, were divided randomly into two groups of 30. They were told only that the experiment was concerned with the way in which first impressions of people are formed.

Stimulus material

Two tape-recordings were made by the same speaker, a woman unknown to the subjects, speaking the same passage but adapting her voice to give an impression of either warmth or coldness.* In the talk, the speaker was describing her job and Asch's six non-central traits were implied as closely as possible, without direct mention. The passage was as follows, with the relevant traits inserted in brackets:

'Good afternoon. I have been asked to speak to you today about my work as a veterinary surgeon and to give you some idea of the variety and interest which I derive from my large and far-flung country practice in North Wales. I qualified as a vet. after training at Liverpool University and I was fortunate to complete my very long training in the minimum time. (INTELLIGENT.) I then obtained a post as assistant to a vet. who was already in established practice and who gave me every opportunity for making use of my newly-acquired professional skills. (SKILFUL.) My work is virtually of two kinds; firstly with small animals, mostly pets, who are brought to morning and evening surgery by their owners for treatment of major or minor ailments or accidents; secondly it is with larger animals of the scattered farms in the valleys and hills around the town itself. In addition to tending sick animals in this way, I am responsible to the Ministry of Health for the testing of herds of cattle for tuberculosis. This has to be done at regular intervals and has become a part of almost every day's work, after morning surgery, and is something with which my assistant is, as yet, unqualified to help me.'

'My day begins with the arrival of the postman at half-past six and, very often, with the ringing of the telephone, as farmers are early risers and crises so often happen during the night, (INDUSTRIOUS.) Surgery over by ten o'clock, I set out in my car for the neighbouring and distant farms. The problems of weather, transport, distances, bad roads and difficult hills are very great in a country practice, but my assistant and I have managed not to let such difficulties deter us from our purpose of reaching a sick animal at all costs. (DETERMINED.)

* Thanks are due to Miss Muriel Judd, Senior Lecturer in Speech and Drama at the College, for invaluable training and assistance in making the recordings.

And because of these hazards of mountain roads and unforeseen weather changes, I have found it essential to learn how to service my own car and keep it in running repair. All too often in the early days of my work here, I found myself having to change a wheel or dismantle the carburettor. (PRACTICAL.) Some parts of the countryside are very lonely, so I always carry a crowbar in my car. I wondered at first whether a gun would be a better weapon of defence, if put to the test, but the crowbar is easy to wield and it gives one a feeling of security. (CAUTIOUS.)

'Long distance visits are usually completed by mid-afternoon, and then I return home to deal with telephone calls, messages and emergency visits to surgery. Sometimes I have to make further local visits. Evening surgery is rarely over by seven o'clock and even then I have my records to complete after supper. But I recommend this work as being interesting, varied and worth while, and ideal for those who enjoy life in the open air and the country, and who are not afraid of hard work.'

Procedure

The experimental groups each heard one of these recordings and they were then asked to write free character sketches of the person. Finally, they were given a checklist like that used by Asch (see Table 3), except that the words 'warm' and 'cold' were added, and they were asked to underline the word of each pair that best fitted the person described.

RESULTS AND DISCUSSION

All six non-central traits were mentioned in a number of the free character sketches in each group, at least by a very close synonym, suggesting that the passage was adequate in conveying the characteristics intended. All 30 subjects who heard the 'warm' recording underlined 'warm' on the checklist, whereas only 18 of the 'cold' group underlined 'cold'.

Considering only these 48 subjects, who did receive the impression intended, their selections among the remaining pairs of traits in the checklist agreed to a significant degree with those marked in Experiment 2 below (in which no actual person was used, reproducing, for the most part, Asch's original experiment). Comparing the warm groups of the two experiments, by ranking the eighteen characteristics according to number of selections, ρ is 0.92; for the two cold groups, it is 0.63.

There are discrepancies, however, between the two experiments, the most marked of which are shown in Table 1. These suggest a greater attribution of

Table 1. Discrepancies between the cold groups of Experiments 1 and 2

Traits	Nos. selecting them (N = 30)	
	Experiment 1	Experiment 2
generous	7	24
ruthless	16	0
self-centred	21	7
humorous	4	17
sociable	5	22
popular	5	23
important	20	29
restrained	28	14

'warm' qualities (e.g., generous, humorous, sociable) and a lesser attribution of 'cold' qualities (ruthless, self-centred, restrained) on the part of the cold group in

the first experiment compared with the cold group of the second. These results should be considered together with the fact that 12 of the cold group did not identify the speaker as cold. There are at least two possible reasons for this. The acting may have been inadequate, and a naturally warm voice not sufficiently concealed: however, 60 per cent of those who were intended to hear a cold voice did, in fact, do so. The more likely explanation is that the occupation described, that of veterinary surgeon, was even more potent in conjuring up expectations than the coldness of the voice: Asch himself (1946) had different results on his checklist according to whether the original traits were attributed to a man, a woman or a child. Certain social and occupational roles are readily associated with warmth from the nature of the relationships or work involved, and an occupation devoted to the nurturance of animals is likely to be among these. The cold group in this experiment was, therefore, confronted with incompatible information, as in the experiments of Pepitone & Hayden (1955) and Haire & Grunes (1950). Some dealt with the situation by ignoring the less definite information.

Centrality

In the case of the 80 per cent of subjects who did not receive the intended impression, the subsequent selection of traits was very similar to that of Experiment 2, where the descriptions 'warm' and 'cold' were made explicitly. This is substantial justification of the attempt to represent these characteristics by manner only. While the efficacy of other traits has not been tested in the same fashion, it is difficult to imagine the use of this experimental method for the presentation of other pairs of qualities demonstrated by Wishner as being central in his sense—'humane' and 'ruthless', for example. The warm-cold dimension is a pervasive one—i.e. it can be expressed by voice, gesture and posture as well as in the content of speech—and this makes it likely to be an important dimension in impression-formation. And, with its connotations of acceptance and rejection, it will be for many people a highly relevant dimension on which to classify people being met for the first time. For some, however, an occupational or similar social classification gives information even more relevant to their expectation systems.

Asch's claim that the warm-cold variable occupied a central position in impression-formation rested on two findings. One of these will be discussed later: that when the subjects were asked to mark the given traits according to their importance in impression-formation, warm and cold were ranked high. More important was the effectiveness of these two traits, as distinct from others Asch tried, in giving differential results on the checklist. This is where Wishner has produced contrary evidence. Claiming that Asch's differential results would be entirely predictable from independent measures of the relationship between the so-called central and other traits, he was able to substantiate this, with a high degree of success, by first asking student subjects to assess their University instructors on relevant semantic differential scales, which afforded the independent measures of relationship required. This alone, of course, does not weaken Asch's claim that warm-cold is a central variable. But Wishner then applied the procedure to make another pair of traits equally 'central', in the sense of differentiating on a checklist, when this new

checklist consisted of traits found independently to be related to the new central variable.

While there is no need to dispute Wishner's main contention that traits interact in predictable ways, which Asch appears to have denied, there are points that should be made about Wishner's experiments. First, his demonstrations are not complete. He has failed to test also whether the warm-cold variable would be effective with the checklist specially composed for the 'humane-ruthless' variable. But, more important, the independent measures of the relation between traits were obtained from the assessments of one occupational group only—University teachers—and by subjects by whom they were already known. Asch's experiments invited the subjects to 'create' persons not already known, and the experiment described above demanded impressions of someone being met for the first time. In first encounters the warm-cold variable may be very important, as argued above, while it may become far less important in someone familiar and with whom one has a formalized relationship, as student with teacher. Moreover, in the assessment of certain occupational groups (such as University lecturer compared, say, with veterinary surgeon) this particular variable may be less salient, and the interaction of traits may be different. In fact, Wishner's findings support his theoretical position to a greater extent than might be expected from so biased a sample of persons to be assessed. Some of the negative findings, however, may be explicable by reference to the particular expectations aroused by the social role studied (for example, 'inflexible' having a negligible correlation with 'unintelligent' and 'popular' with the blunt-polite dimension when University teachers are rated; but high degrees of association when a person has to be 'created'), although to pursue the argument thoroughly it would be necessary to know the distributions of assessments on each trait. That there will be systematic differences in the interrelation of traits associated with different occupations has yet to be shown, however.

Although Wishner is probably right when he says 'a trait is central for those traits correlated with it and peripheral for those traits uncorrelated with it' (p. 108), it could still be the case that certain traits will have more traits in all correlated with them, which increases their chances of being 'central' in Wishner's sense.

EXPERIMENT 2

It seems to have escaped notice that in the original experiment the stimulus lists of traits were *read* to the subjects, despite the possibility of unequal emphasis being given, unintentionally, to the seven words. The transition to visual presentation occurred in 1950 (Haire & Grunes; Kelley), but without explicit mention and in association with other procedural modifications. The experiment now to be reported is a repetition of Asch's original experiment but with visual presentation, performed partly to check this point and partly to gather material that could be analysed more closely than is possible from the published accounts of Asch's data.

Subjects

Sixty women students from the same training college were used, these being in the third year.

Procedure

The traits used by Asch were typed on sheets of paper, thirty containing the word 'warm' in the middle of the list, thirty 'cold'. These were interleaved randomly for distribution. The subjects were asked, in both written instructions and verbally, to read them several times, trying to form an impression of the person described, since afterwards they would be asked to write brief character sketches. These were written after one minute of reading time. Meanwhile, checklists of pairs of traits, those used by Asch (see Table 3) were distributed and the subjects were asked to mark those appropriate to the person conceptualized. Finally, they were asked to turn back to the original seven traits and indicate their rank order of importance in forming their impressions.

RESULTS AND DISCUSSION

Influence of the warm-cold variable

The closeness of the results on the checklist to those of Asch can be seen from Table 2, in which Asch's results have been converted into percentages, and only

Table 2. *Comparison of Asch's and present results on the checklist: Experiment 2*

Traits	Warm group		Cold group	
	Asch per cent	Present per cent	Asch per cent	Present per cent
generous	91	93	8	23
wise	65	77	25	23
happy	90	97	34	53
good-natured	94	90	17	23
humorous	77	80	13	13
sociable	91	90	38	16
popular	84	93	28	16
reliable	94	100	99	100
important	88	50	99	66
humane	86	83	31	33
good-looking	77	70	69	46
persistent	100	100	97	100
serious	100	77	99	96
restrained	77	50	89	93
altruistic	69	70	18	10
imaginative	51	63	19	10
strong	98	90	95	83
honest	98	100	94	93

one of each pair of traits is shown. Discrepancies exist, but they are small, and in no case is the warm-cold influence in the opposite direction.

Table 3 summarizes the findings for the present experiment. To what extent can it be said that the giving of warm or cold has modified the personality traits selected? One way of answering this question statistically is by taking out, from each pair, the trait with the higher score for the warm group and comparing these scores with those received from the cold group. On the Wilcoxon Signed-Rank Test, the difference is significant at the 1 per cent level.

Examining the pairs of traits singly, comparisons can be made by the Chi-Square Test, bearing in mind the strictures of Kjeldergaard & Jenkins (1958) against differences at the 5 per cent level with traits which are probably not independent.

Table 3. Selection of traits on checklist: Experiment 2

Trait	Warm per cent	Cold per cent	Trait	Warm per cent	Cold per cent
*generous	93	23	ungenerous	0	56
*shrewd	16	70	wise	77	23
*unhappy	0	26	happy	97	53
*irritable	7	37	good-natured	90	23
*humorous	80	13	humourless	0	73
*sociable	90	16	unsociable	0	70
*popular	93	16	unpopular	0	53
unreliable	0	0	reliable	100	100
important	50	66	insignificant	7	13
*ruthless	7	53	humane	83	33
good-looking	70	46	unattractive	13	36
persistent	100	100	unstable	0	0
frivolous	7	0	serious	77	96
restrained	50	93	talkative	40	3
*self-centred	17	70	altruistic	70	10
*imaginative	63	10	hard-headed	30	70
strong	90	83	weak	3	16
dishonest	0	0	honest	100	93

* Significantly different at the 1 per cent level (χ^2).

For ten of the 18 pairs the proportions of selections from the warm and cold groups differ to a degree significant at the 1 per cent level.

It would seem that the reading of 'warm' or 'cold' in the original list has again quite substantially modified the impression formed. Moreover, as Bruner, Shapiro & Tagiuri (1958) pointed out, the traits selected are not necessarily denoted by the given one: thus, 'wise' and 'imaginative' are markedly associated with warmth, but are certainly not logically implied by it.

Centrality of the warm-cold variable

The ranking of traits for their importance in impression-formation again demonstrated the dominance of warm and cold: 50 per cent of the warm group and 60 per cent of the cold group placed warm or cold in the first or second rank. In speaking of the centrality of these traits, Asch meant, in part, that the peripheral traits would fall into different hierarchical constellations according to the nature of the central trait. This is shown numerically in two ways with the present data. First, while an equal number in each group (14) put warm or cold in the first place, in the warm group the trait receiving the next highest number of 'firsts' is 'intelligent', and in the cold group, it is 'determined'. Second, and more important,

Table 4. Mean ranks of peripheral traits: Experiment 2

Traits	Warm group	Cold group
intelligent	5½	6
skilful	2	2
industrious	7	4
determined	5½	7
practical	4	3
cautious	3	5

it can be seen from Table 4 that the mean rankings of the peripheral traits differ in the two groups: these are calculated from the rankings of the 28 subjects who put warm or cold first. The trait 'industrious', which carries a low mean rank in association with 'warm', illustrates most clearly the point that relative importance can be transformed by the central traits.

Totality of the impression

Luchins (1948) has criticized Asch's claim that in the character sketches an entire person is represented. Unfortunately Asch makes little detailed reference to this material. Luchins himself, on repeating the experiment, found that a high proportion of the character sketches consisted of the trait names, or synonyms for them, strung out as sentences.

The subjects of the present experiment showed little tendency to be restricted to the given traits. One index of this is the occurrence of sentences bearing no direct relationship, apparently, to any of the given traits: sentences such as, 'Someone interested in weird, unorthodox forms of art and music and with high-brow taste for the theatre'. As many as 59 per cent of all the sentences written (from the assessments of two judges working separately) are independent in this sense. Comparing only the subjects who ranked warm and cold in first place, in their assessment of relative importance in impression-formation, the influence of the warm-cold variable is not apparent in the attribution of gender, there being an equal number of males and females described in each group. There is a difference in mention of occupation (and, by implication, of age): occupations are ascribed more often by the cold group, always with the assumption that the person is an older man or woman, holding a position of responsibility usually, such as a business man or a manageress. Finally, an analysis of the descriptions of physical appearance, given by about half the subjects, yields this difference, that whereas the cold people are always tall and thin, the warm ones are as likely to be small and plump.

REFERENCES

ASCH, S. E. (1946). Forming impressions of personality. *J. abnorm. soc. Psychol.* **41**, 258-290.

ASCH, S. E. (1952). *Social Psychology*. New York: Prentice-Hall.

ASCH, S. E. (1958). The metaphor: a psychological enquiry. In *Person Perception and Interpersonal Behaviour*. (Tagiuri, R. & Petrullo, L., eds.). Stanford: Stanford U.P. Pp. 86-94.

BRUNER, J. & TAGIURI, R. (1954). The perception of people. In *Handbook of Social Psychology*. (Lindzey, G., ed.). Cambridge, Mass.: Addison-Wesley. Pp. 634-654.

BRUNER, J., SHAPIRO, D. & TAGIURI, R. (1958). The meaning of traits in isolation and in combination. In *Person Perception and Interpersonal Behaviour*. (Tagiuri, R. & Petrullo, L., eds.). Stanford: Stanford U.P. Pp. 277-288.

COFER, C. N. & DUNN, J. T. (1952). Personality ratings as influenced by verbal stimuli. *J. Pers.* **21**, 223-227.

GOLLIN, E. S. (1958). Organizational characteristics of social judgment: a developmental investigation. *J. Pers.* **26**, 139-154.

HAIRE, M. & GRUNES, W. F. (1950). Perceptual defenses: processes protecting an organised perception of another personality. *Hum. Relat.* **3**, 403-412.

HAYS, W. L. (1958). An approach to the study of trait implication and trait similarity. In *Person Perception and Interpersonal Behaviour*. (Tagiuri, R. & Petrullo, L., eds.). Stanford: Stanford U.P. Pp. 289-299.

KELLEY, H. H. (1950). The warm-cold variable in first impressions of persons. *J. Pers.* **18**, 431-439.

KJELDERGAARD, P. M. & JENKINS, J. J. (1958). Personality ratings as influenced by verbal stimuli—some negative findings. *J. Pers.* **26**, 51-60.

LUCHINS, A. S. (1948). Forming impressions of personality: a critique. *J. abnorm. soc. Psychol.* **43**, 318-325.

MENSH, I. N. & WISHNER, J. (1947). Asch on 'Forming impressions of personality', further evidence. *J. Pers.*, **16**, 188-191.

PEPITONE, A. & HAYDEN, R. (1955). Some evidence for conflict resolution in impression formation. *J. abnorm. soc. Psychol.* **51**, 302-307.

SHAPIRO, D. & TAGIURI, R. (1958). Some effects of response control on trait impression. *J. Pers.* **26**, 42-50.

WISHNER, J. (1960). Re-analysis of 'Impressions of Personality'. *Psychol. Rev.* **67**, 96-112.

Manuscript received 15 January, 1962

The Measurement of Religious Attitudes in a University Population

By PAMELA K. POPPLETON AND G. W. PILKINGTON

Department of Education, University of Sheffield, and Department of Psychology, University of Sheffield

A scale was constructed for measuring the religious attitudes of students at a British University. Scores obtained from administration of the scale to a representative sample of 500 students show that there is a significant decline in religious belief among students in the Faculties of Arts and Pure Science during early years at the University, but that this decline continues only for the Science students, especially those who later engage in research. Medical students in their final years and students training for the teaching profession, especially the arts graduates, have the highest scores of all. Between denominations it was found that Catholics and members of small evangelical sects scored higher than adherents of other denominations, and between sexes the familiar finding that women tend to be more religious than men was replicated. Some implications of these findings are discussed.

A. INTRODUCTION

The publication in the early 1930's of the Thurstone scales for the measurement of attitudes towards the church, the reality of God, and the influence of religion on conduct stimulated some enquiry into the distribution of these attitudes among different sections of the American population. The subjects most frequently studied were American college students, and their religious behaviour and beliefs have been investigated in both the pre-war era (Carlson, 1934) and post-war (Allport, Gillespie & Young, 1948). But no such information is available about the religious attitudes and affiliations of British students, even though recent years have seen some speculation about a revival of religion in British universities. When the authors wished to discover the direction of recent trends, they felt that it would be worthwhile also to construct a scale which, in addition to answering the questions posed in this enquiry, would be a reliable and valid instrument for use in further investigations. No scale of this nature exists which is adequate for such a purpose. A survey of the literature yields the following summary.

Table 1. *Scales used in studies of religious attitude*

Author	Subject	Designed for	Nationality
1 Thurstone, L. L. & Chave, E. J. (1929)	Attitude towards the Church. The reality of God. Influence on Conduct	General population, but have been used with students	American
2 Stone, S. (1936)	Religiosity Scale	Parents and children	American
3 Brown, D. G. & Lowe, W. L. (1951)	Inventory of Religious Belief	Students	American
4 Ferguson, L. W. (1944)	'Religionism' scale. Derived from Thurstone scales above	Students	American
5 Moreton, F. (1944)	Likert scale	General population	British
6 Eysenck, H. J. (1947)	Social Attitudes scale	Adult students	British

Of the scales mentioned above, (3) was the most relevant to our enquiry, but the statements in it were deliberately chosen so as to encourage outright acceptance or rejection for the purpose of obtaining two well-defined groups. Of the British scales, (6) is a composite scale containing only a few items which concern religion, and (5) was designed for use with the general population. The items which this and some of the other scales contain are rather crude. We felt that expressions of belief amongst students might well be of a more sophisticated nature, and therefore that items should be derived from amongst the student population.

Our interest lay in measuring attitudes towards the Christian religion. This is not because we regard other religious beliefs as unimportant, but to have included them in our scale would have added yet another 'dimension' to a set of beliefs which we suspected were very far from being homogeneous. Statements were invited from a large number of students and also from members of the teaching staff. The first part of the article is concerned with an account of the compilation of the scale, the methodological problems which arose during this stage and the scale's administration to a representative sample of students. The findings are then presented and discussed.

B. METHODOLOGY

It is customary nowadays for the would-be compiler of an attitude scale to be eclectic in his choice of method and to borrow impartially from the many sources that are available. In making his choice he seeks to attain high reliability and validity, and is also influenced by any tactical considerations which may be important for the particular attitude concerned. The paths, however, are well worn, and the procedures followed in this case were:

1. The Thurstone method for the compilation and scaling of items.
2. The Likert method for item analysis.
3. Guttman scale analysis for demonstration of unidimensionality.

Construction of the scale

Using the Thurstone method, two parallel forms of a scale were constructed for a pilot survey. Each consisted of 22 items selected by the usual techniques from an original 156, which referred to many aspects of religious belief and were drawn from student and staff members of the University. These two forms were then given to a group of 121 people from the same population who were asked to respond to each item in one of five categories, 'Strongly Agree', 'Agree', 'Uncertain', 'Disagree' and 'Strongly Disagree'. Half the group had Form A first, and half Form B, the other form being given after an interval of three weeks. From the two sets of scores data were obtained about the reliability of the scale, and some interesting findings emerged concerning different methods of scoring. These, however, will be presented and discussed in a further paper. The method finally adopted was one suggested by Guilford (1954) which uses empirically derived weights for the response categories. The means and standard deviations of the resulting distributions were calculated, and also the product moment correlation between the scores on the two forms. These are shown in Table 2 below.

Table 2. *Statistics of the two parallel forms used in the Pilot Survey*

	Mean	S.D.	Range	Correlation (r)
Form A	87.2	20.5	76.0	} 0.93
Form B	89.5	19.6	74.0	

An item analysis was conducted on both forms of the scale. The method used was that suggested by Likert (1932), as it is computationally simple and gives satisfactory results.



This analysis resulted in the discarding of 23 items, the remaining 21 then being assembled into a final Form C.

One trend in evidence throughout was the gradual elimination of statements which expressed a subtle or sophisticated attitude. The statements which remained were straightforward, and sometimes rather naïve, expressions of belief and disbelief. This may seem surprising in view of the educational level of the population for which the scale was designed. But items in attitude scales are as much an artifact of the processes of statistical analysis to which they are subjected as they are expressions of the beliefs of a given population.

The main sample

As it was intended to examine the differences in religious attitude between faculties and between years at the university, a proportionate, stratified random sample was taken of the whole student population. The total number of students on the University register in February, 1961, was approximately 3000. In order to obtain a sample of 500, every sixth name was taken at random from each year within each faculty. Each of these 500 students was asked to complete Form C as well as a short questionnaire concerning his religious beliefs and practices (see appendix).

It was thought to be essential for any enquiry of this nature that anonymity should be guaranteed. A code number system was, therefore, adopted which indicated faculty and year blocks whilst retaining anonymity for the individual, and care was taken to explain this system in an accompanying letter. The questionnaires were circulated by post, a stamped, addressed envelope being enclosed for reply. Response was excellent, 463 of the 500 students approached returning completed questionnaires, which gives a response rate of 92.6 per cent. Included among the non-respondents are seven forms which were either returned uncompleted or which were not completed according to the instructions. Another six were erroneously sent to non-Christian students and are excluded from the sample. Among the respondents seven apparently regarded the presence of numbers on their forms as compromising their anonymity and had deleted them. These forms could not, therefore, be classified according to faculty or year of study and hence were not included in those analyses.

The final Form C

The administration of this form differed in only one respect from the previous forms. Six response categories were provided instead of five. The 'uncertain' category was replaced by columns for 'mildly agree' and 'mildly disagree'. Opinion on the question of forced choice differs, and in some respects it was an unsatisfactory step for a scale of this kind. Many respondents, when returning the forms, indicated their wish to express uncertainty in response to some items. As far as the scoring of the scale was concerned, however, both these categories received the same weight and, in fact, the scalogram analysis of the form was made slightly easier and more accurate. Nevertheless, for future use it is probably advisable to retain the uncertain category (see appendix). The forms were scored using the empirically derived weights obtained previously. Figure 1 shows the distribution of these scores.

The distribution is markedly bimodal in form, which is consistent with a tendency that other workers have noticed. Thouless (1935), for example, suggests that true doubt and scepticism are rare in religious beliefs, and Carlson (1934) also found a similar trend in attitudes towards the reality of God. The neutral zone, which is cross-hatched in the diagram, was determined by locating the zero point of the scale. For this, the method described by Stouffer *et al.* (1950) as the 'fold-over' method was used. Different individuals tend to make use of one or another category in their responses to items, and an 'Emphasis score' can be calculated for each person by scoring 2 points for each 'Strongly Agree' or 'Strongly Disagree', 1 point for each 'Agree' or 'Disagree' and 0 for each 'Mildly Agree' or 'Mildly Disagree'. The relationship between the E score and the subject's total score can then be shown in a scatter diagram, the lowest point of which indicates that score at which intensity of response is minimal. Such a scatter diagram was constructed and resulted in the U-shaped form which is typically found with a bi-polar variable. The zero point was located round about the score value of 85, just below the median.

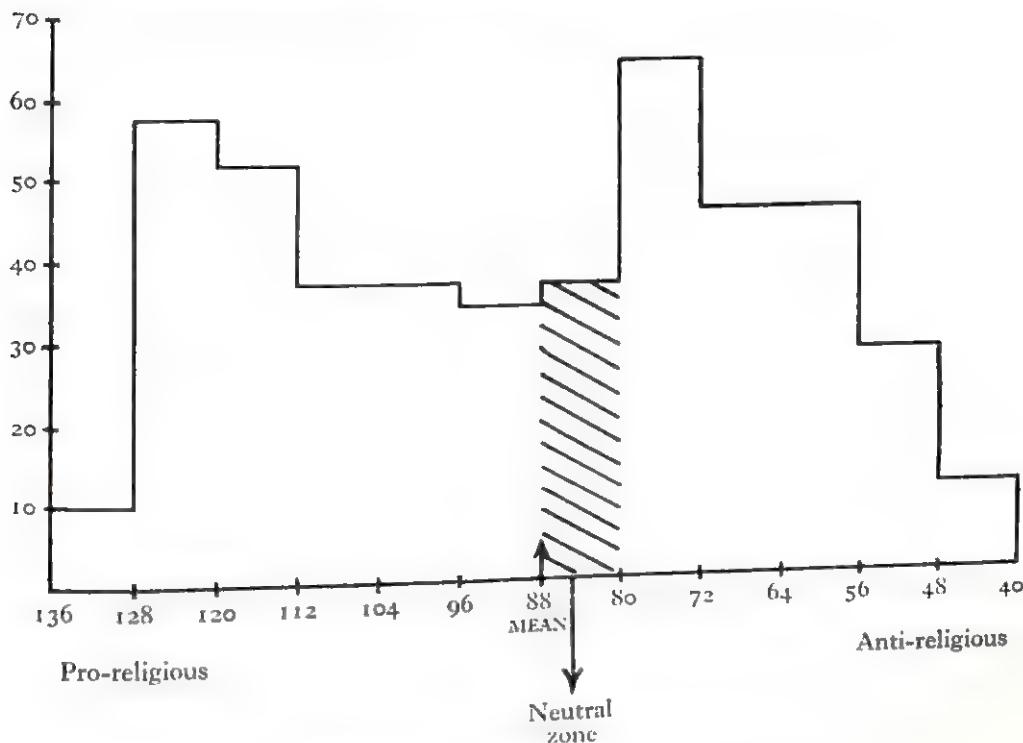


Fig. 1. Distribution of total scores. Form C. N = 463

Reliability

For the purpose of calculating the reliability of the new scale, it was split into three so as to test all parts. Scores were calculated for all subjects in three sub-sets of seven items each; intercorrelations between these parts were calculated and corrected for length (r_{12} , r_{13} , r_{23}). The results are shown in Table 3.

Table 3. Reliability data

Sub-set	1	2	3
Mean	29.1	29.8	30.4
S.D.	8.0	8.7	8.2
r_{12}	0.95		
r_{13}		0.96	
r_{23}			0.97

Reliability was calculated by substituting values of the variances in Cronbach's (1947) formula for coefficient alpha. This gave a value of $\alpha = 0.97$, a very high measure of reliability.

Unidimensionality of the scale

A scale may be reliable without being homogeneous in the sense that all items contribute equally to the variance. A commonly used test of homogeneity is Guttman's scalogram analysis. Accordingly, such an analysis was conducted using the Cornell Technique described by Guttman (1947). Responses to each item were analysed from one hundred completed forms which were chosen so that the scores fell into a normal distribution. The criteria laid down by Guttman for calculating a coefficient of reproducibility were met, and an overall reproducibility of 87.76 per cent was achieved. This does not quite reach Guttman's requirement of 90 per cent but is nevertheless very high. It appears likely that this

is the result of the item analysis carried out in the pre-test with the parallel forms. The reproducibility of the items is shown in the appendix.

Validity of the scale

One of the main problems was that of finding behavioural indices against which to validate the scale. The only factual information available which could be checked independently of the respondents' reports was that concerning membership of student religious groups, e.g. The Student Christian Movement, the Methodist and Catholic societies. Recourse, therefore, was inevitably had to respondents' statements about their religious activities and beliefs. An additional sheet attached to the questionnaire asked for such information (see appendix), and from it a markedly pro-religious and a markedly anti-religious group were selected in the pilot survey. A *t*-test between the mean scores of these groups on Form B showed them to be significantly different at $p < 0.01$. Even this preliminary version of the scale showed some evidence of validity. A similar kind of procedure was used to determine the validity of Form C as administered to the main sample of respondents. Because of the bimodal nature of the distribution non-parametric statistics were used, median scores for a pro-religious and anti-religious group being worked out.

The pro-religious group (Group A) consisted of 107 respondents who reported (a) active membership of a church, (b) church attendance of three times or more during February, 1961, and (c) saying private prayers at least once weekly. The anti-religious group (Group B) comprised 109 respondents who described themselves as either atheists or agnostics. Group A had a median score of 116, and Group B a median score of 60. There was no overlap at all between the scores of these two groups. Moreover, the 54 members of Group A who report saying private prayers daily and who are members of a student religious society scored even higher: their median score was 122. It is evident that the scores obtained on the scale corresponded in a consistent way with other indices of religious behaviour and belief. It may be concluded that Form C is a reasonably sensitive and accurate instrument which can help to provide answers to some of the questions which are posed in this enquiry.

Estimating bias in the sample

When a postal questionnaire is used for the collection of survey material difficulties arise from the fact that response is never one hundred per cent (see Bell, 1961). Some people respond at once, others need one or two reminders, others do not respond at all. If these non-respondents differ in important ways from the rest, then the sample is biased, since those who have replied cannot be regarded as representative of the whole group. Hence every effort was made to ensure that response was as complete as possible, short of actually visiting students, by sending out reminder letters. The numbers of respondents to successive letters and of the various kinds of non-respondents is shown in Table 4.

Table 4. *Analysis of successive responses*

Respondents	
First	350
Second	71
Third	42
Unsatisfactory	7
Non-Christian (omitted from all analyses)	6
Non-respondents	24
Total	500

To determine whether or not the group of non-respondents differs in any of their characteristics from the rest of the sample, the method suggested by Zimmer (1956) was used. This involves the estimation of trends from the initial response group through the second and third respondents, so that any factor which varies systematically with the probability of response can be regarded as a characteristic of the non-response group. Two kinds of factors are present: (1) Objective information regarding faculty, year of study and

denomination. (2) Information derived from the indices of religious behaviour already used, i.e. active church membership, church attendance, the saying of private prayers, the holding of religious belief and membership of some student religious group.

It is arguable that the probability of non-response will depend at least as much upon the meaning to the subject of the topic being investigated. For example, a firm believer may well be more ready to answer a questionnaire of this kind than someone who is merely apathetic about religious matters. An analysis was therefore undertaken to determine whether or not a trend of this nature was apparent.

1. Group differences in objective characteristics

For the purpose of this analysis, second and third respondents were grouped together and compared with first and non-respondents. These are called Medium, High and Low groups respectively. Table 5 shows the significance of the differences between them. Chi-square was calculated in each case, and where an overall significant difference was found which suggested that they were not likely to be random samples from a common population, they were compared two at a time.

Table 5. Differences between respondent groups

Characteristic	d.f.	Overall			High v. Low			High v. Med.			Med. v. Low		
		χ^2	p	d.f.	χ^2	p	d.f.	χ^2	p	d.f.	χ^2	p	
1 Faculty	6	13.8	<0.05	3	3.77	>0.20	3	16.7	<0.001	3	0.24	>0.950	
2 Year	6	10.8	>0.05	—	—	—	—	—	—	—	—	—	
3 Sex	2	2.0	>0.30	—	—	—	—	—	—	—	—	—	

N.B. Where the null hypothesis can be rejected, p values are given for one-tailed tests of significance.

The only real difference which emerges here is that between faculties, and it is the high and medium responding groups that differ. Initially, the proportion of first respondents was higher in the faculties of Arts and Pure Science (81 per cent and 79 per cent) as against Medicine and the Applied Sciences (64 per cent and 68 per cent). For these differences, χ^2 reached a value of 12.61, which is significant at $p < 0.001$. The balance was redressed to some extent in response to subsequent letters, but an overall difference remained. One possible explanation on this effect may be a factor peculiar to this University, where the Arts and Pure Science faculties are in much closer physical proximity to the centre of university life than are Medicine and the Applied Sciences.

Although no significant differences are apparent between the years in high and low response, they do vary in their initial response to the questionnaire. Education and research students, all of whom are postgraduate, show a higher rate of initial response than the rest of the sample ($\chi^2 = 4.9, p < 0.05$). If they are separated from the remainder of the students who are classified as fourth year and who belong to Medicine and the Applied Sciences, then the latter show a lower rate of response than the rest ($\chi^2 = 8.8, p < 0.01$).

2. Group differences in religious behaviour

In this case, trends were estimated from the characteristics of the first, second and third responding groups, and the same statistical procedure applied as previously. The findings are presented in Table 6.

The only statistically significant difference emerges between initial and final response concerning active church membership. The percentage figures do, however, reveal some suggestive tendencies, and these are shown in Figure 2 on the following page.

It may, therefore, be concluded that, in objective characteristics, the sample is slightly underrepresented in the faculties of Medicine and the Applied Sciences, and that, in religious characteristics, readiness to respond is associated only with reported active church membership; but that there is suggestive evidence that activity on two of the other indices of religious behaviour also gradually declines over the three response groups in the sample. Thus it is possible that the results to be reported in the next section very slightly over-estimate the degree of religious activity and belief of the Sheffield student population.

Table 6. *Differences between respondent groups in religious behaviour*

Characteristic	Overall			1st v. 3rd			1st v. 2nd			2nd v. 3rd		
	d.f.	χ^2	p	d.f.	χ^2	p	d.f.	χ^2	p	d.f.	χ^2	p
1 Active church member	2	7.3	<0.05	1	4.2	<0.05	1	0.4	>0.50	1	1.4	>0.20
2 Church attendance	2	2.2	>0.30									
3 Group member	2	2.2	>0.30									
4 Prayer saying	2	0.9	>0.50									
5 Holding belief	2	0.7	>0.70									
6 Denomination	4	2.3	>0.50									

N.B. Where the null hypothesis can be rejected, p values are given for one-tailed tests of significance.

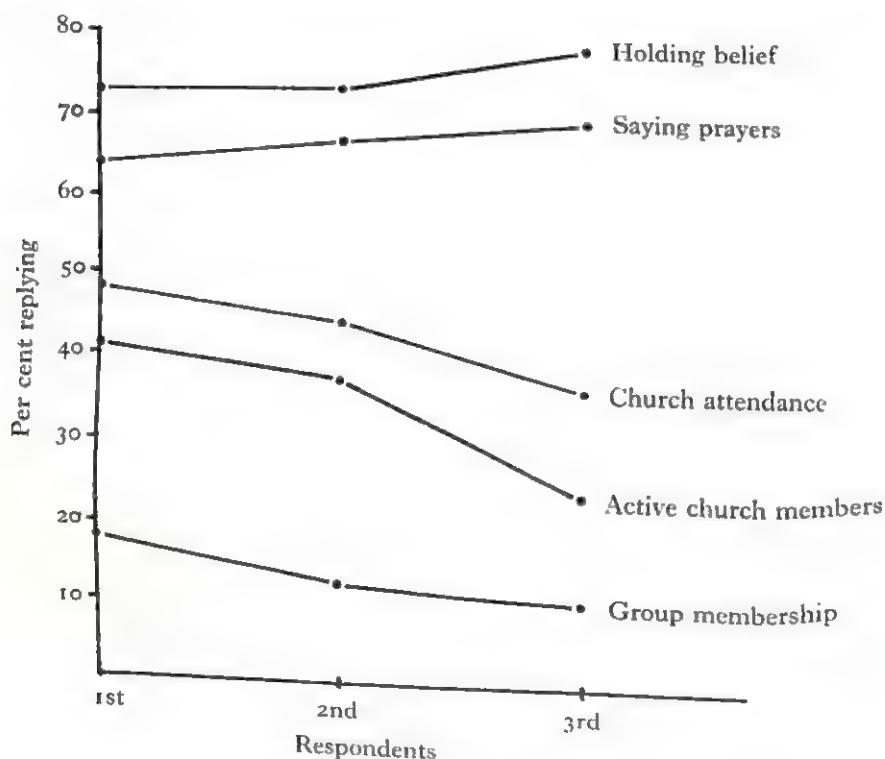


Fig. 2. Proportions of respondent groups on each of five indices

C. RESULTS

I. Analysis of trends for faculties and years

Table 7 below shows the median scores for faculties and years of study. It is based on scores from the religious attitude scale, Form C. The total of 456 is less than that of the whole sample by seven, this being the number of respondents whose forms could not be replaced in year and faculty groups because of deleted code numbers.

Table 7. Median scores for faculties and years of study

	N	1st yr.	2nd yr.	3rd yr.	4th yr.	Total
Arts (including Architecture, Law and Sociology)	157	83.0	74.5	95.5	104.5	84.5
Pure Science	141	100.0	79.0	74.0	78.5	81.0
Medicine and Dentistry	67	86.0	105.0	79.5	112.0	91.5
Applied Sciences	91	88.0	95.0	89.0	82.5	90.0
Total	456	89.0	85.5	83.5	90.5	87.0

Because of the bimodal nature of the score distributions, the above data were analysed by the median test extended for k independent groups described by Siegel (1956, p. 179). Although no overall statistically significant differences were found, some suggestive trends are apparent, and more detailed analysis of the data provides some interesting results.

Table 8. Analysis of median scores for faculties and years

	d.f.	χ^2	p
(a) Between faculties			
Total	3	1.68	>0.50
First year	3	1.87	>0.50
Second year	3	6.77	>0.05
Third year	3	5.15	>0.10
Fourth year	3	5.33	>0.10
(b) Between years			
Total	3	2.02	>0.50
Arts	3	4.08	>0.20
Pure Science	3	4.42	>0.20
Medicine and Dentistry	3	6.95	>0.05
Applied Science	3	2.10	>0.50

N.B. p values are given for two-tailed tests of significance.

(a) *Differences between faculties.* The difference between faculties almost reaches significance in the second year. This point is marked by a large drop in the scores of arts and pure science students as compared with the others. By the fourth year, however, the arts students' median score has risen almost to the level of that of the medical students (see Table 7). Taken together, fourth year students in the faculties of Medicine and Arts score significantly higher than their contemporaries in the faculties of Pure and Applied Science ($\chi^2 = 4.5$, for 1 d.f., $p < 0.05$).

(b) *Differences between years.* Reference to Table 7 shows that in the faculty of Pure Science third year students score much lower than those in the first year, and this difference is significant ($\chi^2 = 3.82$, for 1 d.f., $p < 0.05$). The decline is apparent by the second year in both Arts and Pure Science and if the smaller faculties of Architecture, Law and Sociology are excluded then this decline is also significant ($\chi^2 = 3.3$, for 1 d.f., $p < 0.05$).

In the Medical faculty the overall differences between the years almost reach significance, and a test of the difference in scores between the third and fourth years gives $\chi^2 = 4.2$ (for 1 d.f., $p < 0.05$). It should, however, be noticed that the differentiation into years of study of the medical students is slightly different from that used for the remainder of the sample. Most medical students go straight from school into the second year of their course; such students have here been counted

as first year. Similarly, fifth and sixth year medical students have been treated so as to correspond to the fourth year students of other faculties, who mainly consist of research students and of postgraduates in the Education Department. An analysis of the scores of these two groups shows that the education students scored significantly higher than the research students ($\chi^2 = 2.71$, for 1 d.f., $p < 0.05$). The arts and pure science students within the Education Department also exhibited a wide difference in their median scores. As numbers for calculating χ^2 here were too small, the exact probability associated with such a difference in score was worked out; this proved to be $p = 0.056$. The scores are shown in Table 9 below.

Table 9. *Median scores for 4th year education and research students*

	Arts	Pure Science	Applied Science	Total
(N = 26)				
Education students	113.5	91.0	—	106.5
(N = 35)				
Research (postgraduate)	—	72.0	81.0	79.0

2. Analysis of denominational differences

Table 10 shows the median scores of respondents according to their religious affiliation. They are divided into two groups: (a) those who report that they are active church members of the denominations shown, and (b) those who say that they are not active church members, but who were nevertheless brought up in some religious denomination. A third small group consists of 14 students who had no denominational upbringing, but since none of these report active church membership, they are included in group (b) above. The heading 'small evangelical sects' comprises such groups as Christadelphians, Plymouth Brethren, Jehovah's Witnesses, Church of Nazarene, and Church of the Latter Day Saints.

Table 10. *Median scores of denominations*

Denomination	N	Median score	Q
(a) <i>Active church members</i>			
Roman Catholic	24	122.0	5.3
Small evangelical sects	11	118.0	*
Methodist	25	114.5	10.8
Congregational and Baptist	18	114.5	6.0
Church of England	103	109.5	12.1
Total	181	113.6	10.0
(b) <i>Not active church members</i>			
Roman Catholic	4	*	*
Small evangelical sects	10	72.0	*
Methodist	37	70.0	13.3
Congregational and Baptist	13	76.5	*
Church of England	201	75.8	15.5
No denominational upbringing	14	85.0	*
Mixed denominational upbringing	3	*	*
Total	282	75.0	13.6

* Sample too small to give reliable scores.

Median tests of significance between the active members of denominations were carried out, and the results are shown in Table 11 on following page.

Table 11. Relationships between denominations

	Values of χ^2			
	Roman Catholic	Methodist	Congregational and Baptist	Church of England
Roman Catholic	—	—	—	—
Methodist	4.15*	—	—	—
Congregational and Baptist	2.93	.10	—	—
Church of England	12.81†	.51	1.35	—

* Significant at $p < 0.05$ (one-tailed test).

† Significant at $p < 0.001$ (one-tailed test).

3. Analysis of sex differences

Table 12 shows differences in median scores between men and women.

Table 12. Differences between sexes

(a) Total sample			
Sex	N	Mdn	Q
Men	357	83.3	21.6
Women	106	104.5	21.0
Total	463	88.8	21.4

(b) Active and non-active church members			
Sex	Active		Non-active
	N	Mdn	Q
Men	122	111.7	10.2
Women	59	115.7	8.4
	235	74.3	12.2
	47	77.9	14.9

A median test between men and women in the whole sample gives $\chi^2 = 6.2$ which for 1 d.f. is significant at $p < 0.01$. Within the active and non-active categories, however, significant differences between the median scores of the two sexes emerged only for the non-active group. For active men versus active women $\chi^2 = 2.3$, which for 1 d.f. means $p > 0.10$; and within the two non-active groups $\chi^2 = 3.6$, which for 1 d.f. means $p < 0.05$. (In all cases one-tailed tests of significance were used.)

4. Results of the questionnaire on religious activities and beliefs

A wealth of interesting information was accumulated as a result of the use of this questionnaire. As we propose to analyse and discuss these data in detail in a separate paper, only the overall percentages of students replying 'yes' and 'no' to the main items are presented here; they are shown in Table 13 below.

Table 13. Percentage responses to questionnaire items

	Yes per cent	No per cent
Hold some form of religious belief	74	26
Say private prayers	65	35
Attended church once or more during the month	46	54
Are active members of a church	39	61
Say private prayers daily	31	69
Attended church four or more times during the month (once a week?)	23	77
Members of a student religious group	16	84

D. DISCUSSION OF RESULTS

1. *Difference between faculties and years of study*

A number of American studies agree in finding a decline in religious beliefs and activities during early years at college, but which is checked or reversed during later years (Arsenian, 1943; Allport, Gillespie & Young, 1948; Telford, 1950; Brown & Lowe, 1951; Argyle, 1957, p. 44). Very little evidence is available, however, concerning such trends in this country, although Furlong (1961) has recently suggested, on the basis of subjective impressions, that 'many students at all the universities appear to turn from faith to unbelief in their first year. Or at least to an abandonment of religious practice'. That there is a general decline in religious beliefs and activities in both countries during the 18-30 age range is better substantiated. Argyle (1958, pp. 65-67) summarizes many studies, but implies that it may well be the factors associated with this time of life rather than age itself that are responsible, e.g. career-making. Thus, in this enquiry, we are to some extent concerned with the influence of career-making, as represented by a course of study followed at a University, upon religious beliefs.

The overall results for years of study, shown in Table 7, are at least consistent with the American findings mentioned above. But none of the differences reported is significant, so that these data can give no direct support to the general hypothesis that a university course of study, of whatever kind, causes an initial turning away from previously held religious convictions. On the other hand, the decline in religious belief from the first to the second year in the faculties of Arts and Pure Science is clear and also statistically significant; and the American studies, which report similar tendencies, appear to have been carried out on students of these kinds rather than on medicals and technologists. Another trend which is particularly obvious in our data is the marked decline in scores over the years among students in the faculty of pure science; this is especially clear if pure science students in the department of education are excluded from the fourth years in this faculty. Our findings here agree with those of Thoday (1957). Working at the University of Birmingham, she noted a steady and significant decline in church attendance among pure science students over their years at University. Taken together, these results may mean that the much popularized science versus religion conflict is still a real one for many science students in the universities.

Is there a relationship between students' religious beliefs and their choice of career? Our data alone can, of course, give no hint as to which operates as the causal factor, but at least three tendencies are clearly in evidence. First, students scores. Secondly, those about to enter the teaching profession, especially the arts graduates, have relatively high scores; a possible interpretation of this finding is that students with strong religious beliefs are more likely to choose teaching as a career than other better paid occupations, but in which they would find less scope for their religious values. Thirdly, medical students in the last two years of their course emerge as highly religious. Furlong (1961) puts forward the interesting view that a movement back to religion occurs amongst medical students during

their years of clinical study, because then 'students turned away from theory to face the sign of human suffering in their daily experience'. Our data confirm that such a trend occurs, although whether this is the result of facing suffering and death or of pressures to conform in a strongly socialized profession, is not clear.

2. Denominational differences

The findings reported under this heading agree with many other studies in showing that Catholics nearly always rank higher than other denominations on various indices of religious belief and activity, e.g. Allport, Gillespie & Young (1948), Telford (1950), Brown & Lowe (1951) and Gilliland (1953). But if an attempt is made to compare our data in more detail with American findings, then the result is somewhat unsatisfactory because the range of denominations and the lines of demarcation between them are so different in the two countries. The nearest comparable investigations are those of Kirkpatrick (1949) and Lawson & Stagner (1954).

Their results, as well as ours, suggest that religious attitudes within the various Protestant denominations are more sharply differentiated, and sometimes more variable, than is often supposed. Thus, to include together as members of one group a number of different Protestant denominations, as has been done in many studies, can obscure important differences. For example, our results show that the small evangelical sects are almost as high in religious conformity as the Catholics. Moreover, among Methodists and adherents of the Church of England, there is a considerable variation in strength of religious belief, as is shown by their dispersion of scores.

As might be expected, the strength of religious belief among those who are no longer active church members is much less than that of the active church members. What is more surprising is that the original denominational affiliation does not influence the degree of later religious belief, as is shown by the small differences in median scores of the various denominations. Another unexpected finding is the relatively high score of those in this group who had no denominational upbringing. Although numbers are very small, and the difference in score does not reach statistical significance, this result does suggest that there is a need to investigate the religious beliefs of this kind of group in a separate study.

3. Sex Differences

The overall results reported here will cause no surprise. Argyle (1958, pp. 71-79) draws attention to a large number of investigations which show consistently that women are more religious than men on many different criteria, and Allport, Gillespie & Young (1948, p. 10) refer to 'this general well-known law of sex-differences'. Only a very few workers, such as Moreton (1944), Brown & Lowe (1951) and Gilliland (1953), have failed to replicate this general finding, and in all these enquiries there is some reason to believe that the samples chosen were more than usually homogeneous in their religious outlook. Our data amply support this 'law', except within the category of active church members.

4. Overall trends in students' religious beliefs

Do our data support the view that there is at present a trend back to religion in the Universities, as has often been asserted recently, for example, by Furlong (1961)? In the absence of scores on the scale from the general population (matched with the students for age and social class) and from former generations of students, no firm answer can be given. The general picture, however, that emerges from this particular sample is one of the acceptance of religious ideas by the majority of students; this is shown by the high percentages reporting some form of religious belief and saying private prayers. Moreover, only during the third undergraduate year does the median score fall below the zero point of the scale. On the other hand, those students whose religious life is at all active constitute only a minority of the sample, as reference to the proportions reported for church attendance, active church membership and the daily saying of private prayers makes clear. It could be argued that these proportions would need to be much higher before it could plausibly be claimed that religious ideas and practices are a really important influence on the lives of most University students.

The authors would like to thank Miss M. Jessop and Miss M. Billett for secretarial assistance; Mr Eagle, Technician, Department of Psychology, for the line drawings and Mrs E. M. Denman, Assistant Registrar, for help in the sampling arrangements.

REFERENCES

ALLPORT, G. W., GILLESPIE, J. M. & YOUNG, J. (1948). The religion of the post-war college student. *J. Psychol.* **25**, 3-33.

ARGYLE, M. (1958). *Religious Behaviour*. London: Routledge & Kegan Paul.

ARSENIAN, S. (1943). Change in evaluative attitudes during four years of college. *J. appl. Psychol.* **27**, 338-349.

BELL, C. R. (1961). Psychological versus sociological variables in studies of volunteer bias in surveys. *J. appl. Psychol.* **45**, 80-5.

BROWN, D. G. & LOWE, W. L. (1951). Religious beliefs and personality characteristics of college students. *J. soc. Psychol.* **33**, 103-129.

CARLSON, H. B. (1934). Attitudes of undergraduate students. *J. soc. Psychol.* **5**, 202-212.

CRONBACH, L. J. (1947). Test 'reliability': its meaning and determination. *Psychometrika*, **12**, 1-16.

EYSENCK, H. J. (1947). Primary social attitudes: I. The Organisation and measurement of social attitudes. *Internat. J. Opin. Attit. Res.* **1**, 49-84.

FERGUSON, L. W. (1944). A revision of the primary social attitude scales. *J. Psychol.* **17**, 229-241.

FURLONG, M. (1961). Religion in the universities. *Sunday Times*, 11th June, 28.

GILLILAND, A. R. (1953). Changes in religious beliefs of college students. *J. soc. Psychol.* **37**, 113-116.

GUILFORD, J. P. (1954). *Psychometric Methods*, Second edition. London: McGraw-Hill.

GUTTMAN, L. (1947). The Cornell technique for scale and intensity analysis. *Educ. Psychol. Mesmt.* **7**, 247-280.

KIRKPATRICK, C. (1949). Religion and humanitarianism: a study of institutional implications. *Psychol. Monogr.* **63**, No. 9 (whole No. 304).

LAWSON, E. D. & STAGNER, R. (1954). The Ferguson 'religionism' scale: a study in validation. *J. soc. Psychol.* **39**, 245-256.

LIKERT, R. (1932). A technique for the measurement of attitudes. *Arch. Psychol.* No. 140.

MORETON, F. E. (1944). Attitudes to religion among adolescents and adults. *Brit. J. educ. Psychol.* **14**, 69-79.

SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioural Sciences*. New York: McGraw Hill.

STONE, S. (1936). A comparative study of the religious attitudes of parents and their children. Unpublished M.A. thesis: Univ. of Minnesota.

STOUFFER, S. A. *et al.* (1950). *Measurement and Prediction*. Princeton: University Press.

TELFORD, C. W. (1950). A study of religious attitudes. *J. soc. Psychol.* 31, 217-230.

THODAY, D. (1957). Halls of residence. *Univ. Quart.* 12, 45-56.

THOULESS, R. H. (1935). The tendency to certainty in religious belief. *Brit. J. Psychol.* 26, 16-31.

THURSTONE, L. L. & CHAVE, E. J. (1929). *The Measurement of Attitude*. University of Chicago Press.

ZIMMER, H. (1956). The validity of extrapolating nonresponse bias from mail questionnaire follow-ups. *J. appl. Psychol.* 40, 117-121.

Manuscript received 7 February 1962

APPENDIX
Religious Attitude Scale

FORM C

Below are 21 statements which concern religious beliefs. Please indicate the extent to which you agree or disagree with each of them. On the right-hand side of the page you will find five alternative answers. Place a cross opposite each statement in the column which best represents your opinion. For example:

Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree

More time in broadcasting should be allotted to agnostic speakers.

Please do not leave out any statements even if you find it difficult to make up your mind.

We should be grateful if you would also fill in answers to the questions on the final sheet. All the information given will be treated as strictly confidential. Thank you for your co-operation.

Rep	Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree
87	6	6	5	4	2
95	2	2	2	5	7
91	2	2	4	6	6
85	2	4	5	6	6

Rep	Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree	
	—	—	—	—	—	
5. Religious faith is merely another name for belief which is contrary to reason. (10.05)	93	2	2	4	5	7
6. The existence of disease, famine and strife in the world makes one doubt some religious doctrines. (7.43)	82	2	2	4	6	6
7. The miracles recorded in the Bible really happened. (1.22)	84	6	6	4	2	2
8. It makes no difference to me whether religious beliefs are true or false. (6.20)	85	3	3	3	4	5
9. Christ atoned for our sins by His sacrifice on the cross. (0.62)	96	7	6	4	2	1
10. The truth of the Bible diminishes with the advance of science. (9.00)	87	2	2	3	6	6
11. Without belief in God life is meaningless. (0.73)	86	7	6	4	2	1
12. The more scientific discoveries are made the more the glory of God is revealed. (1.47)	95	6	6	3	2	2
13. Religious education is essential to preserve the morals of our society. (2.64)	86	6	5	4	2	2
14. The proof that Christ was the Son of God lies in the record of the Gospels. (1.53)	89	6	6	3	2	2
15. The best explanation of miracles is as an exaggeration of ordinary events into myths and legends. (8.71)	93	2	2	4	6	6
16. International peace depends on the world-wide adoption of religion. (2.06)	82	6	6	5	3	2
17. If you lead a good and decent life it is not necessary to go to church. (7.33)	85	2	3	4	6	6

Rep	Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree
18. Parents have a duty to teach elementary Christian truths to their children. (2.70)	86	6	5	3	2
19. There is no survival of any kind after death. (10.37)	85	1	1	2	5
20. The psychiatrist rather than the theologian can best explain the phenomena of religious experience. (8.88)	83	2	2	3	6
21. On the whole, religious beliefs make for better and happier living. (3.32)	88	6	5	3	2

The numbers in brackets after each statement refer to the Thurstone scale values of the items. Values range from 0-11. Low values indicate pro-religious and high values anti-religious attitudes.

The column headed Rep shows the coefficient of reproducibility for each item. Weights are indicated by the numbers in the ruled columns.

Sample Survey Questionnaire

Would you please answer the following additional questions?

- (1) Age (in years)
- (2) Sex: FEMALE or MALE (underline one)
- (3) Would you describe yourself as an active member of a church? YES OR NO (underline one)
- (4) If 'yes' to question (3) to which denomination do you belong?
- (5) If 'no' to question (3), in which denomination were you brought up (if any)?
- (6) How many times did you attend church during the month of February?
- (7) Are you a member of any student religious group (e.g. the S.C.M., Padley Society, Methodist Society, etc.)? YES OR NO (underline one)
- (8) Do you say private prayers? YES OR NO (underline one)
- (9) If so, do you say them:
 - (a) At least once daily? YES NO
 - (b) At least once weekly? YES NO
 - (c) Less frequently? YES NO

(please underline appropriately)

(10) Do you consider yourself to be a holder of some form of religious belief?
YES NO (please underline appropriately)

(11) If 'no' to the previous question, which of the following categories best describes your beliefs?

- (a) Agnostic. YES NO (please underline appropriately)
- (b) Atheist. YES NO (please underline appropriately)
- (c) Other (please describe as briefly as you can)

(12) Please add here any other information or comments you consider useful or relevant to the survey.

The Response and Self-Generated Behaviour of Severely Disturbed Children and Severely Subnormal Controls

BY BEATE HERMELIN AND N. O'CONNOR

*Medical Research Council, Social Psychiatry Research Unit,
Institute of Psychiatry, London, S.E.5*

Children admitted to an institution for severely subnormal patients and subsequently diagnosed as schizophrenic were compared with a group of children of the same age, sex and I.Q. in the same hospital, who were free from schizophrenic characteristics. The two groups showed a similar amount of exploration of and orientation towards stimuli. Though the experimental group gave the same amount of social response and approach behaviour as the controls, the former also tended to retreat from the experimenter at times. Significantly more children talked in the control than in the experimental group.

Activity of a non-object orientated kind was more frequently observed in the psychotic children than in the controls. While the amount of orienting behaviour varied according to the various stimulus situations in both groups, activity which did not occur in response to external stimuli did not.

INTRODUCTION

A recent annotation in the *Lancet* (1961) discussed the general uncertainty about the nature and aetiological basis of a schizophrenic syndrome in children. This uncertainty is due in part to the lack of a well differentiated personality in the young child, which makes features of personality disintegration hard to define. Children with different aetiology and varying degree and nature of organic damage may show some or all of the types of behaviour which are regarded as characteristic of schizophrenic children. In addition the existence of schizophrenia in children seems by some definitions to rest on the assumption that the condition is a purely functional one, and that its onset follows a clearly established period of normal development. Bourne (1955) argues that severe behaviour disturbance as well as severe subnormality can result from gross parental deprivation. However, such deprivation cannot be established in all cases of disturbed imbeciles. In addition, present methods of neurological and electro-encephalographic investigations seem not yet sufficiently refined to rule out organic involvement with certainty. At least in the case of idiots and imbeciles the occurrence of psychotic or autistic symptoms is complicated by the presence of extensive lesions. Thus, Earl's (1934) study of catatonic psychosis in idiocy suggests neurological involvement. He regarded the symptoms as 'a form of schizophrenia played out on the psychomotor level rather than the symbolic'.

Mayer-Gross, Slater & Roth (1954) have defined childhood schizophrenia or autism as 'independence of an unresponsiveness to the environment'. This definition forms the basis of the present study, which investigated the behaviour of severely disturbed subnormal children and controls in various stimulus situations.

Selection of subjects

The present sample comprises severely subnormal children in an institution, some of whom show severe behaviour disturbances while others do not. From a list of children with a history of behaviour disturbance, aged between 5 and 16 years, those who had severe motor or sensory handicaps were excluded, as were also those for whom no I.Q.s were available, or where I.Q.s of 25 or below had been obtained. A psychiatrist who knew the children well was then asked to judge which of the children would be diagnosed as schizophrenic. The diagnosis, which was based on the observed presence of manneristic or autistic behaviour, was accepted regardless of whether it was made as 'primary schizophrenia' or as a superimposed condition on mental deficiency. Subsequently the psychiatrist was asked to estimate the patients' symptomatology according to the nine criteria of childhood schizophrenia described by Creak *et al.* (1961).

Twelve patients were selected, diagnosed as schizophrenic, and manifesting at least three of the nine criteria. The average number of criteria present was five. Eleven of the 12 children were regarded as showing 'gross and sustained impairment of emotional relationship with people', a characteristic regarded by Creak *et al.* as the most crucial in the schizophrenic syndrome. Five out of the 12 children were judged by the psychiatrist to possess 'islets of normal, near normal or exceptional intellectual function or skill against a background of serious retardation'. This characteristic was regarded by some of the members of Creak's working party as a *sine qua non* for childhood psychosis.

The control group was matched with the experimental group for sex, age and I.Q., although the significance of intelligence test scores at this level of functioning is somewhat equivocal. According to psychiatric diagnosis and rating, none of the controls showed schizophrenic characteristics and only two scored on one of the nine points mentioned; in both cases this was 'gross retardation of speech'. We debated whether children with conditions such as tuberous sclerosis and phenylketonuria should be excluded from the sample. However, as such conditions are frequent in the mentally deficient, it was decided that both experimental and control groups should reflect this fact.

Data regarding information about a history of birth trauma or relevant disease in early infancy was obtained for each subject, as was also information about neurological examinations. The presence of epilepsy and gross cerebral lesions associated with, for example, meningitis or cerebral palsy, was also noted. This information is summarized in Table 1. Only 2 children in each group showed no obvious signs of brain damage or birth trauma.

Table 1. *Number of relevant clinical signs*

	Epileptic seizures	Abnormal birth conditions	Gross cerebral lesions	Strabismus
Psychotics	3	5	3	3
Controls	0	8	3	4

METHOD

Each child was observed individually for six sessions of 15 minutes each. Observations were carried out through a one-way screen from an adjoining room. Sessions for each child took place on three consecutive days, one occurring in the morning and another in the afternoon. Experimental sessions occurred in the same order for each subject. Randomization of conditions was not attempted, because it was considered that the content of the session became progressively more complex.

The children were placed in a large 24 ft by 12 ft room, which for the first session was empty except for such fixtures as lights, light switches, windows, radiators and doors. After the children had thus become familiar with the room, exactly the same procedure took place in the second session. This in turn was followed by four further 15-minute periods of observation, in which visual, auditory, manipulative and social stimulation were provided in this order.

For visual stimulation three 1 ft by 1 ft 6 in. displays were fixed to the wall at a height of 7 ft. One was a plain piece of red paper, the second a wallpaper pattern of shapes, and the third a representational picture showing people and animals. Two auditory stimuli were introduced in the auditory session, each for a duration of 3 minutes. They were white noise and a nursery tune, emitted from two loudspeakers on opposite walls. Sound intensity at the source was 70 db against a room background noise of 40 db. Each auditory session consisted of 3 minutes without sound, followed by 3 minutes of white noise or music, then 3 minutes silence, 3 minutes of music or white noise, and then finally 3 minutes without sound. For the manipulative session three toys were provided, a piece of string, a large top and a transparent U-shaped maze puzzle containing a sweet.

In the social session one of the experimenters entered the room after 3 minutes, bringing a chair in with her. She sat still and silent at the far end of the room for 3 minutes, behaving as passively as possible. In the subsequent 3-minute period the experimenter approached the child, attempting to establish some contact by either cuddling or tickling or by some form of play. In the fourth 3-minute period the experimenter spoke to the child with the aim of investigating his understanding of language by asking him to carry out simple commands, and tried to involve the child in conversation and make him answer simple questions. In the fifth 3-minute period the child was again left alone in the room. There were thus five 3-minute periods in each social session, the first and the last being periods in which the experimenter was not present.

Recording of data

Twenty-one items of behaviour were listed and their occurrence in any 1-minute period checked off on a scoring sheet. The list of items is given in Table 2.

Table 2. *Items scored*

1. Smiling	12. Lying down (sitting)
2. Crying	13. Covering eyes or ears
3. Vocalizing	14. Looking at pictures
4. Grimacing	15. Turning towards voice
5. Rocking	16. Playing with toys
6. Twisting	17. Approaching person
7. Finger play	18. Retreating from or ignoring person
8. Door	19. Responding to person (non-verbal)
9. Windows	20. Carrying out verbal commands
10. Light switches	21. Talking to experimenter
11. Walking about (running)	

Each item of behaviour occurring within a period of 1 minute was only scored once, so that the maximum score for each item on each occasion was 15. Many forms of behaviour were scored simultaneously within each 1-minute period. In addition, after each minute the position of the child was marked on a scale drawing of the room. Lines dividing the floor

space into thirty-two equal areas 1 yard square were represented on this plan. Thus a mobility score for each child at sessions 1-5 was obtained. Due to one of the experimenters not being available for recording, no mobility scores were obtained in the social session.

RESULTS

Two kinds of basic data were used for analysis. Where possible, a frequency score for an item was used. However, when the distribution of this departed too far from normality, the number of children in either group who scored on any item rather than the frequency score of that item was compared. Following a technique similar to that of Birch (personal communication) and Thomas & Chess (1957), single items were combined to form categories. Thus frequency scores were cross summed over items and compared between the two groups by analysis of variance. Frequency scores of single items were subsequently compared by *t* tests.

The first analysis of variance compared behaviour items of the two groups in each of the sessions, which could reasonably be interpreted as responses to external stimuli, with behaviour items which could not easily be so described. The latter, called provisionally self-generated behaviour, included items 1-7 of Table 2, i.e. making hand and finger movements, rocking, twisting, making noises, and grimacing, smiling and crying when alone. Looking out of the window, switching the light on and off, and trying the door handle were regarded as response behaviour. So were looking at the visual displays, turning towards the noise sources, playing with the toys, and responding to the experimenter in the social session. Such items were: 8, 9, 10 and 14-21 from Table 2. The results of the analysis of variance comparing these behaviour categories over groups and sessions are shown in Table 3.

Table 3. *Analysis of variance between groups, categories and sessions*

Source	SS	Df	MS	F	P
Between G					
Between C	10.134	1	10.134	> 1.0	N.S.
Between S	598.072	1	598.012	83.826	> 0.001
C × G	150.433	5	30.086	4.217	> 0.01
G × S	85.578	1	85.578	11.996	> 0.001
C × S	32.574	5	6.515	> 1.0	N.S.
C × G × S	136.946	5	27.389	3.839	> 0.01
Between people within groups	11.776	5	2.355	> 1.0	N.S.
Residual	3377.260	22	153.512		
Total	7890.522	1106	7.134		
	19334.12	1151			

As can be seen from the table, there is a significant interaction between categories and sessions, which shows that, while response behaviour varies from session to session, the amount of self-generated behaviour remains relatively constant in both groups. The significant interaction of categories and groups demonstrates that, while the groups do not differ significantly in the total amount of response behaviour, the experimental subjects show more self-generated behaviour than the controls. This is illustrated in Figure 1. Because of the significance of the interactions, main terms are not discussed.

In the experiment, response behaviour falls into two classes. One is behaviour directed towards the exploration of permanent features in the experimental room, such as door, windows and lights. We called this 'exploratory behaviour'. The other is orienting responses towards stimuli which were introduced at one session only, such as pictures, noises and toys, i.e. 'orienting behaviour'. An analysis was therefore carried out, comparing the amount of exploratory and orienting behaviour

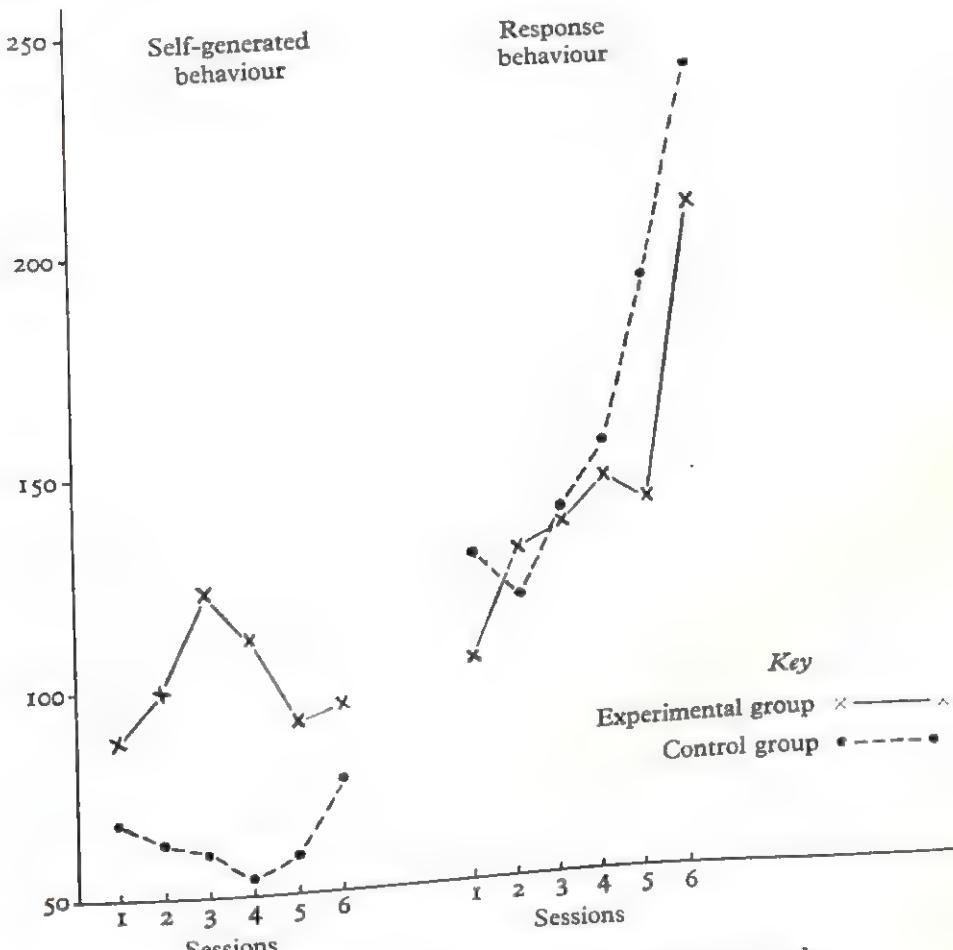


Fig. 1. Frequency mean scores for psychotics and controls.

in the two groups. There was no significant difference, nor did any of the interactions reach a level of statistical significance. However, comparing the groups on item 16, 'response to toys', showed that the controls gave significantly more manipulative responses than the disturbed children ($p = 0.05$).

Turning to the social behaviour of the subjects, the items 17-21, i.e. approaching, responding to or retreating from the experimenter, carrying out verbal commands and answering questions, were combined and the two groups compared. This showed that the controls gave significantly more social responses than the experimental subjects ($p = 0.01$). The next step was to compare separately the items of

behaviour which had been combined in the social responses category. Owing to their skewed distribution these were tested by X^2 . There was no significant difference in the number of children from either group who made non-verbal approaches or responded non-verbally to the experimenter. Only 1 out of 12 children in each group failed to score on these items. However, 10 children in the experimental group also retreated temporarily from the experimenter or ignored her at times compared with 5 in the controls. This significant difference ($X^2 = 4.4$; $p = 0.05$) indicates that more experimental subjects than controls showed an alternating approach-retreat reaction to a person.

Avoiding non-social stimuli, measured by covering eyes and ears, did not differentiate the groups, and only very few in either group scored on these items. There was no difference in the extent to which the children in either group carried out verbal commands. There was, however, a significant difference when the number who talked in either group was compared. Nine out of 12 controls and only 4 out of 12 experimental subjects responded verbally ($X^2 = 3.84$; $p = 0.05$). On the other hand, the extent of non-verbal vocalization, i.e. singing, grunting, squealing, etc. did not differentiate the groups.

Finally, comparing the mobility records of the children, it was found that there was no difference between the groups nor between the sessions in the amount of walking or in the area covered.

DISCUSSION

As we noted, this study was based on Mayer-Gross, Slater & Roth's (1954) statement that schizophrenic children showed independence of and unresponsiveness to the environment. For the group of children tested here this statement needs modification. They were not unresponsive to environmental stimuli whether permanently present or temporarily introduced. To both these kinds of stimuli they responded as much as did the control group, except in the sessions where toys were provided. This difference might be accounted for by the difference between the groups in the amount of 'self-generated' activity. The schizophrenics showed much more of this than did the controls, and the items in which this difference was most marked were rocking and hand and finger movements. These movements may have interfered with the manipulation of objects, while they left orientation to visual, auditory or social stimuli relatively little affected. As far as social behaviour is concerned, in addition to an alternating approach-retreat behaviour which is marked in the schizophrenic children, the main difference is one of frequency of speech. This absence of speech cannot be accounted for by a complete inability to vocalize, as non-verbal vocalization does not differ between the groups. Thus while, apart from speech behaviour, this study does not support a conclusion that sub-normal schizophrenic children are less responsive to their environment than sub-normal controls, it indicates that the former show more behaviour which is not orientated towards external stimuli than do the latter. The frequency with which such behaviour occurred remained relatively constant, and was independent of the specific contents of the various sessions.

The approach-retreat behaviour of the psychotic children in the social session may be explicable in terms of social 'withdrawal' or in terms of restlessness or fluctuation of attention. If restlessness were involved, one might have expected more overall motor activity in the psychotics than in the controls. This was not the case.

Finally, it should be remembered that in this experiment order of conditions was kept constant so that sequential effects cannot be isolated. Accommodation to the experimental setting may have played a part in the increasing of response behaviour over sessions in both groups.

We are grateful to Dr B. H. Kirman and the Staff of the Fountain Hospital for providing facilities for this research. We would also like to thank Dr A. E. Maxwell and Dr P. H. Venables for statistical advice.

REFERENCES

ANNOTATIONS (1961). Fact and fiction in infant rearing. *Lancet*, ii, 813-814.

BOURNE, H. (1955). Protophrenia: a study of perverted rearing and mental dwarfism. *Lancet*, ii, 1156-1163.

CREAK, Mildred, *et al.* (1961). Schizophrenic syndrome in children. *Lancet*, ii, 818.

EARL, C. J. C. (1934). The primitive catatonic psychosis of idiocy. *Brit. J. med. Psychol.*, 14, 230-253.

MAYER-GROSS, W., SLATER, E. & ROTH, M. (1954). *Clinical Psychiatry*. London: Cassell. (Reprinted 1955.)

THOMAS, A. & CHESS, S. (1957). An approach to the study of sources of individual differences in child behaviour. *J. clin. exp. Psychopath.*, 18, 347.

Manuscript received 26 May 1962

A Note on the Cultural Aspects of the WAIS Vocabulary Subtest in Relation to British Mental Patients

By A. ABI RAFI*

St. Andrew's Hospital, Thorpe, Norwich

SUMMARY

An attempt has been made to study Robertson and Batcheldor's rank order of difficulty of the WAIS vocabulary subtest on a fresh sample of British mental patients. The results give general support to their findings with minor divergences, particularly in relation to the items *slice* and *perimeter*. It is suggested that the present order of difficulty of items would be preferable to Wechsler's, especially when dealing with psychiatric subjects. Furthermore, it is more advantageous to apply this test in full when dealing with such cases rather than to follow Wechsler's 'discontinue' direction.

PROCEDURE

Robertson & Batcheldor (1956) found that certain items in Wechsler's (1955) order of presentation are misplaced in terms of their relative difficulty for British mental patients. Robertson & Batcheldor's ranking, and their item *florin* as a substitute for *penny*, were adopted and the rearranged vocabulary items were administered to 63 male and 38 female chronic schizophrenics. The age distribution was as follows: 20-34 years, males 32, females 20; 35-44 years, males 15, females 7; 45-59 years, males 16, females 11.

RESULTS

Table 1 shows the number of patients giving fully correct, partly correct, and incorrect answers to each item, the numbers omitting it, the mean score per item and the rank order of items according to mean score. The last two columns show Robertson & Batcheldor's ranking by mean scores and Wechsler's order of presentation of items.

As can be seen from Table 1, the results support Robertson & Batcheldor's (1956) findings regarding the order of difficulty of items. The two samples of the British mental hospital population, from two areas separated geographically by slightly over 100 miles, show a fairly high measure of agreement as to the relative difficulty of the WAIS vocabulary subtest items. Items 1 to 10 appear roughly in that order of difficulty in the two samples. Item 21 appears to be less difficult, so is item 36, in the present sample. The differences (up to \pm 3 levels) in the rank order of difficulty of the remaining items are not very large and probably do not warrant changing their placement.

The author's thanks are due to Dr W. J. McCulley, Medical Superintendent, St Andrew's Hospital, for facilities to carry out this investigation.

* Now Consultant Clinical Psychologist, Aley, Lebanon.

Table 1. Rearranged order of presentation of vocabulary test

Words	Fully correct answers	Partly correct answers	Incorrect answers	Omitted answers	Mean score per item	Rank of mean score	Robertson's rank of mean score	Wechsler's order of items
SHIP	99	1	1	0	1.99	1st	1st Equal	2nd
BED	97	3	1	0	1.97	2nd	3rd	1st
FLORIN	97	0	2	2	1.94	3rd	1st Equal	3rd
HASTEN	94	1	1	5	1.89	4th	4th	12th
COMMENCE	90	2	4	5	1.82	5th	5th	15th
REPAIR	77	7	15	2	1.61	7th	6th	5th
BREAKFAST	77	24	0	0	1.78	6th	7th	6th
ASSEMBLE	73	7	11	10	1.53	8th	8th	9th
CONCEAL	72	3	16	10	1.47	9th	9th	10th
DOMESTIC	64	14	13	10	1.42	10th	10th	19th
					Equal			
ENORMOUS	31	58	4	8	1.20	13th	11th	11th
WINTER	42	58	1	0	1.42	10th	12th	4th
					Equal			
PONDER	52	17	16	16	1.21	12th	13th Equal	16th
CONSUME	53	13	18	17	1.19	14th	13th Equal	20th
FABRIC	38	31	15	17	1.07	16th	15th	7th
CAVERN	44	11	18	28	0.99	18th	16th Equal	17th
OBSTRUCT	44	16	26	15	1.04	17th	16th Equal	22nd
REGULATE	29	24	36	12	0.82	19th	18th	14th
TERMINATE	36	7	18	40	0.79	20th	19th	21st
TRANQUIL	31	7	14	49	0.69	22nd	20th	29th
SLICE	16	77	5	3	1.09	15th	21st	8th
MATCHLESS	31	3	36	31	0.65	24th	22nd	25th
					Equal			
SENTENCE	9	56	26	10	0.74	21st	23rd	13th
RELUCTANT	28	12	29	32	0.68	23rd	24th	26th
CALAMITY	19	27	29	26	0.65	24th	25th	27th
					Equal			
SANCTUARY	18	15	37	31	0.51	28th	26th	24th
					Equal			
FORTITUDE	24	6	23	48	0.54	27th	27th	28th
REMORSE	13	22	25	41	0.48	30th	28th	23rd
COMPASSION	23	13	28	37	0.59	26th	29th	31st
ENCUMBER	25	1	17	58	0.51	28th	30th	37th
					Equal			
DESIGNATE	10	4	24	63	0.24	34th	31st	18th
AUDACIOUS	9	12	17	63	0.30	33rd	32nd	34th
TANGIBLE	6	5	28	62	0.17	36th	33rd Equal	32nd
IMPALE	14	8	9	70	0.36	32nd	33rd Equal	39th
OMINOUS	5	6	26	64	0.16	37th	35th	35th
PERIMETER	16	10	17	58	0.42	31st	36th	33rd
EDIFICE	9	2	14	76	0.20	35th	37th	30th
TIRADE	3	6	13	79	0.12	38th	38th	36th
TRAVESTY	3	1	12	85	0.07	39th	39th	40th
PLAGIARIZE	0	2	10	89	0.02	40th	40th	38th

REFERENCES

ROBERTSON, J. P. S. & BATCHELDR, K. J. (1956). Cultural aspects of the Wechsler Adult Intelligence Scale in relation to British mental patients. *J. Ment. Sci.* 102, 612-618.

WECHSLER, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: The Psychological Corporation.

Manuscript received 28 May 1962

On the Dual Nature of Extraversion

BY S. B. G. EYSENCK AND H. J. EYSENCK*

Institute of Psychiatry, Maudsley Hospital

A factorial study is reported of a 70-item matrix containing extraversion-introversion, neuroticism, and lie scale items; the correlations were obtained from a sample of 300 men and women, and questionnaire responses were also available from another similar sample of identical size for the purpose of confirming certain findings. The questions to be answered related (1) to the unitary nature of extraversion, and (2) to the independence of extraversion from adjustment. A third problem raised was the possibility that there might be two identifiable components of extraversion, sociability and impulsiveness; the study was designed to investigate the existence and relationship of these two traits. The results showed (1) that extraversion may be regarded as a unitary factor, depending somewhat on the definition of the term 'unitary'; (2) that extraversion and adjustment are essentially independent; and (3) that sociability and impulsiveness do emerge as separate traits, correlating about 0.5 with each other in two independent samples. It was also found (4) that *sociability* has a slightly positive correlation with adjustment, whilst *impulsiveness* has a slight negative correlation.

INTRODUCTION

In her excellent review of 'extraversion-introversion as a dimension of personality', Carrigan (1960) raises the question of the *unidimensionality* of extraversion. Her conclusion is that 'the unidimensionality of extraversion-introversion has not been conclusively demonstrated' (p. 355); she further points out that several joint analyses of the Guilford and Cattell questionnaires show 'that at least *two* independent factors are required to account for the intercorrelations between the E-I variables' (*ibid.*). The nature of these two factors is suggested by a quotation from Mann (1958), who in his discussion of his own results suggests the possibility that 'Factor III corresponds to the American conception of extroversion, with its emphasis on sociability and ease in interpersonal relations, while Factor IV corresponds to the European conception of extroversion, with its emphasis on impulsiveness and weak super-ego controls' (p. 108). This distinction already appears in the Guilfords' (1934) paper on 'Factors in a typical test of introversion-extraversion', where they try to account for the failure of American researchers to validate McDougall's predictions by arguing that while American workers have been largely concerned with the 'social factor' which appears in their analysis as factor *a*, McDougall has been largely concerned with 'impulsiveness' or factor *c* in their analysis. The factorial study here reported was designed to throw some light on this problem.

Another problem also raised by Carrigan is the relation between extraversion and adjustment; her final comment is that 'a clear-cut answer cannot be given' (p. 357). She also hypothesizes, however, that 'a good case can be made for identifying

* We are indebted to the Research Fund of the Maudsley and Bethlem Royal Hospitals for the support of this investigation.

Social Extraversion as a factor of "well-adjusted" extraversion' (p. 338). Lack of self-control, on the other hand, may reflect 'maladjusted extraversion' (p. 339). Here, also, a study designed to investigate this possibility would seem to be timely.

THE EXPERIMENT

A factor analysis, using the method of principal components, was carried out on a matrix containing 66 extraversion-introversion (E) and neuroticism (N) items; the items used are given in Table 1. Also included was a score from a Lie scale containing 18 items: Sex, Age (dichotomized) and number of '?' responses. The items were selected from a consideration of previous analyses (Eysenck, 1956b; 1959; Eysenck & Eysenck, in 1962 press) and from a review of the literature, in the expectation that they would be likely to throw light on our two questions. All 70 items were intercorrelated by means of product-moment correlations, and four factors extracted. These were then rotated, retaining orthogonality, into a close approximation to Thurstone's simple structure solution, graphical methods being used; a discussion of this rotation is given later on. Three hundred subjects in all were used, of whom 133 were male and 167 female; the mean age of the sample was 27.73.

The sample consisted of evening-class students, housewives, and subjects contacted by members of the Department in various ways to take part in divers experiments; many of them formed part of a panel of subjects maintained in the Department to safeguard the supply of subjects. The intelligence and social status of the group as a whole would be somewhat above the average, and it could not be said to be in any sense of the word a random sample of the population; nevertheless it would probably be true to say that it resembled such a sample better than the usual student population taking part in questionnaire studies of this type. It is perhaps relevant that in previous studies samples similarly collected did not differ significantly from the population norms on the Maudsley Personality Inventory (Eysenck, 1959), suggesting that for our purposes the method of sampling may not have been inappropriate.

Factor loadings on the first four factors are given in Table 1; the fourth factor has high loadings on only two items (70: 0.618; and 8: 0.617) and is clearly a doublet, as both of these are concerned with playing practical jokes on other people. Factor 1 is an E factor, Factor 2 an N factor; identification in terms of item content is clear and in line with previous analyses and identifications (Eysenck, 1956b; Eysenck & Eysenck, in 1962 press). It is the third factor on which interest hinges, and for the purpose of discussion this factor has been plotted against Factor 1 in Fig. 1. For the sake of clarity, items scored in the introverted direction have been reversed; these items have an R after their number. Paying attention only to items having saturations of 0.25 or more on E, we can see that these split up very evenly into two groups which have been indicated on the diagram by enclosing lines. One group contains exclusively items having reference to Sociability (S), the other a little less exclusively items having reference to Impulsiveness (Imp.). Items typically defining the former factor are: Keeps in background on social occasions (R), Quiet in social group (R), Reserved and distant (R), Limits acquaintances to a select few (R), Cracks jokes and tells stories to friends, Takes initiative in making friends, Hates introducing people to each other (R), Likes to mix socially with people, etc. Typical items defining the latter factor are: Thinks things over before acting (R), Acts on impulse, Acts on spur of the moment, Easy-going person, Prefers action to planning for action, Tends towards reckless optimism, Would do almost anything for a dare, Happy-go-lucky, Plans ahead (R), etc. The S group is perhaps more homogeneous than the I group of items, and it is possible that a more appropriate name for the latter group might be found; for the purpose of this paper the term 'impulsiveness' will be used without prejudice to alternative interpretations. The two aspects of this bi-polar factor are seen to correspond well with the two types of extraversion postulated by the Guilfords (1934) and Mann (1958).

As an independent check on two of the points raised, the questionnaire was administered to another similar sample of 300 subjects, and scores obtained for Neuroticism, using the standard procedure of the MPI for this purpose (Eysenck, 1959). (All the MPI questions had been included in the present questionnaire.) Also scored were Sociability and Impulsiveness, using in each case the most highly loaded fourteen questions on that factor (i.e. those with the highest +ve and -ve loadings; these two groups are, by accident, well

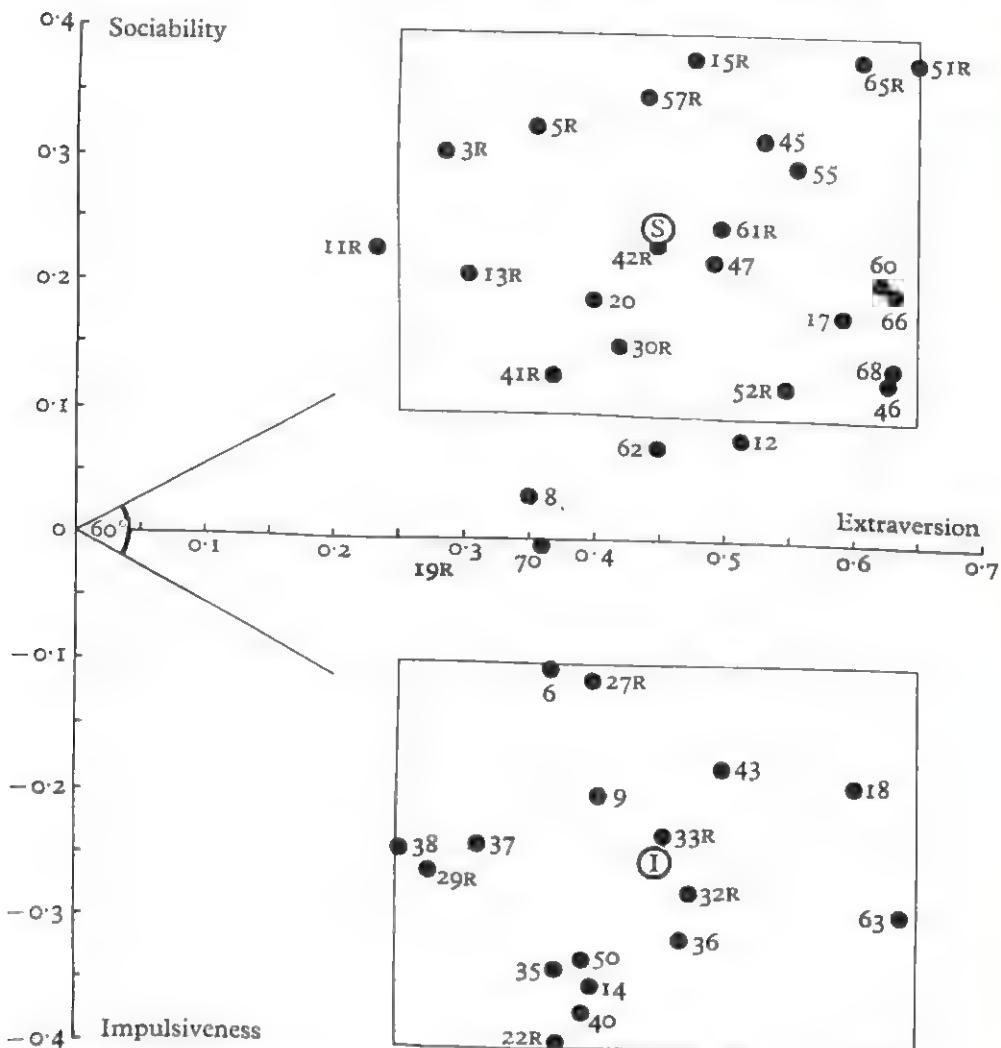


Fig. 1. Plot of high-loading items on Factors I (Extraversion) and III, showing break-up of E-items into Sociability and Impulsiveness.

matched for saturations on the E factor). The correlation between S and Imp. was 0.468 ($p < 0.01$); N correlated -0.133 with S and 0.116 with Imp. (both significant at the 0.05 level), suggesting that these factors, too, were reasonably but not quite independent of abnormality, as indeed had been suggested by their loadings on N in the original factor analysis. N and E (i.e. S + Imp.) correlated -0.011 ; in other words, the two factor scores are quite independent. We thus find that S correlates negatively with abnormality and I positively, thus leaving a zero correlation between N and E, where E = S + Imp. In geometric terms, imagine N to protrude from the plane of the paper in Fig. 1 at right angles; the plane of the paper could be tilted so as to raise Imp. and lower S in the third dimension, without affecting the orthogonality between N and E.

Table I

	E	N	Sv. Imp.	IV
1. Do you sometimes say the first thing that comes into your head?	0.211	0.279	0.173	-0.233
2. Can you usually solve a problem better by studying it alone than by discussing it with others?	-0.153	-0.050	-0.011	-0.079
3. In a group, do you hate having to introduce people to each other?	-0.288	-0.113	-0.309	0.101
4. Do you very much enjoy good food?	0.207	-0.048	-0.095	-0.008
5. Are you a person who is not much given to cracking jokes and telling stories to your friends?	-0.356	-0.016	-0.328	-0.112
6. Do you often crave excitement?	0.369	0.323	-0.103	0.342
7. Do you drink alcohol only in moderation usually?	0.018	-0.253	-0.068	-0.097
8. Do you enjoy practical jokes?	0.349	-0.069	0.035	0.617
9. Would you rate yourself as an impulsive individual?	0.406	0.422	-0.200	-0.305
10. When you are drawn into a quarrel, do you prefer to 'have it out' to being silent hoping things will blow over?	0.196	-0.016	0.124	0.011
11. Do you mind selling things, or soliciting funds for a cause in which you are interested?	-0.233	0.057	-0.228	-0.047
12. Do you like to be in a situation with plenty of excitement and bustle?	0.516	0.070	0.085	0.145
13. If you want to learn about something, would you rather do it by reading a book on the subject than by discussion?	-0.304	0.022	-0.210	-0.228
14. Do you often act on the spur of the moment without stopping to think?	0.400	0.372	-0.349	-0.232
15. Are you reserved and distant except to intimate friends?	-0.477	-0.117	-0.379	0.211
16. When the odds are against your succeeding in some enterprise, do you think it worth while to take a chance?	0.230	-0.006	-0.156	0.111
17. Can you readily get some life into a rather dull party?	0.591	0.113	0.184	-0.079
18. Are you ordinarily a carefree individual?	0.603	-0.218	-0.194	0.156
19. Do you like working alone?	-0.295	-0.037	0.014	-0.259
20. Do you enjoy opportunities for conversation so that you rarely miss a chance of talking to a stranger?	0.403	0.138	0.196	-0.070
21. Do you have difficulty in falling asleep easily at bedtime?	-0.159	0.380	0.081	0.054
22. Are you inclined to stop and think things over before acting?	-0.372	-0.311	0.396	0.221
23. Lie scale	0.177	0.277	-0.102	-0.029
24. ? responses on questionnaire	0.085	-0.031	-0.160	0.130
25. Age	0.069	0.222	0.092	0.347
26. Sex	-0.007	-0.163	-0.001	0.393
27. Do you tend to be slow and deliberate in movement?	-0.400	0.051	0.112	0.359
28. Do you often need cheerful, sympathetic company to 'cheer you up'?	-0.057	0.644	0.042	0.115
29. Do you feel it essential to plan ahead carefully before beginning any undertaking?	-0.276	-0.132	0.258	-0.072
30. On the whole, do you prefer the company of books to people?	-0.420	0.030	-0.153	-0.186

Table I—*continued*

	E	N	S v. Imp.	IV
31. If you are annoyed by something, do you find it absolutely necessary to talk to somebody to 'let off steam'?				
32. Is your motto to take matters of everyday life with proper seriousness rather than to 'laugh and be merry'?	0.201	0.290	0.027	-0.181
33. Do you tend towards an over-cautious pessimism?	-0.475	0.068	0.274	-0.162
34. Do you often have a restless feeling that you want something but do not know what?	-0.454	0.281	0.227	-0.030
35. Would you describe yourself as an easy-going person not concerned to be precise?	0.100	0.537	-0.066	0.111
36. Do you tend towards a rather reckless optimism?	0.373	0.005	-0.335	0.042
37. Would you do almost anything for a dare?	0.468	0.115	-0.313	0.088
38. When people shout at you, do you shout back?	0.313	-0.138	-0.236	0.234
39. Other things being equal, would you prefer the job of a farmer to that of a life insurance salesman?	0.254	0.189	-0.239	0.065
40. Are you given to acting on impulses of the moment which later land you in difficulties?	-0.200	-0.027	-0.022	0.045
41. Would you rather spend an evening by yourself than go to a dull party?	0.392	0.419	-0.372	-0.102
42. Does your natural reserve generally stand in your way when you want to start a conversation with an attractive stranger of the opposite sex?	-0.368	-0.122	-0.130	-0.261
43. Are you happiest, when you get involved in some project that calls for rapid action?	-0.450	-0.008	-0.235	0.196
44. Do you sometimes feel happy, sometimes depressed, without any apparent reason?	0.500	-0.102	-0.181	-0.049
45. Do you usually take the initiative in making new friends?	0.075	0.541	-0.139	-0.029
46. Would you rate yourself as a lively individual?	0.531	0.049	0.320	-0.048
47. Would you be very unhappy if you were prevented from making numerous social contacts?	0.627	-0.062	0.129	-0.088
48. Are you inclined to be moody?	0.490	0.180	0.223	0.053
49. Do you have frequent ups and downs in mood, either with or without apparent cause?	-0.197	0.668	0.022	0.149
50. Do you prefer action to planning for action?	-0.055	0.717	0.086	0.048
51. Are you inclined to keep in the background on social occasions?	0.393	-0.060	-0.327	0.036
52. Is it difficult to 'lose yourself' even at a lively party?	-0.653	0.000	-0.382	0.167
53. Do you ever feel 'just miserable' for no good reason at all?	-0.548	0.045	-0.124	0.114
54. Do you often find that you have made up your mind too late?	-0.087	0.653	-0.045	-0.187
55. Do you like to mix socially with people?	-0.048	0.326	-0.081	0.267
56. Have you often lost sleep over your worries?	0.557	-0.080	0.298	-0.014
57. Are you inclined to limit your acquaintances to a select few?	-0.210	0.482	0.029	-0.010
58. Are you often troubled about feelings of guilt?	-0.442	0.022	-0.353	-0.029
59. Are your feelings rather easily hurt?	-0.161	0.452	-0.032	0.060
60. Do you like to have many social engagements?	-0.106	0.378	-0.100	-0.176
61. Are you inclined to be shy in the presence of the opposite sex?	0.623	0.116	0.210	0.021
62. Do you nearly always have a 'ready answer' for remarks directed at you?	-0.499	-0.023	-0.249	0.259
63. Would you rate yourself as a happy-go-lucky individual?	0.448	-0.070	0.074	-0.175
	0.638	-0.066	-0.292	0.158

Table 1—continued

	E	N	S v. Imp.	IV
64. Have you often felt listless and tired for no good reason?	-0.087	0.532	-0.136	0.036
65. Are you inclined to keep quiet when out in a social group?	-0.609	-0.014	-0.380	0.129
66. Can you usually let yourself go and have a hilariously good time at a gay party?	0.633	0.029	0.200	0.016
67. Do you like work that requires considerable attention?	-0.128	-0.173	0.058	-0.051
68. Do other people regard you as a lively individual?	0.634	-0.042	0.141	-0.161
69. Do you often feel disgruntled?	-0.093	0.632	0.042	0.177
70. Do you like to play pranks upon others?	0.359	-0.023	-0.044	0.618

In all this discussion we have, of course, assumed the validity of the questionnaire responses. While there is independent evidence in favour of this assumption (S. B. G. Eysenck, 1962; Eysenck & Eysenck, 1962) it would certainly be desirable to establish some of the conclusions here reached on a firmer basis by using more objective methods of measurement of the variables involved. Some such measures have been used in other researches (Eysenck, 1957, 1960b; Eysenck & Claridge, 1962) to establish the independence of N and E. S could possibly be measured by using sociometric devices (Eysenck, 1956a) and Imp. might be measured along the lines suggested by Singer *et al.*, (1956) and by Cohen (1959). The present study, while suggestive, is by no means decisive.

DISCUSSION

It is clear that there are (at least) two statistical models which can adequately represent the facts as depicted in Fig. 1. We can say that the results demonstrate the existence of a factor of extraversion which is made up (*inter alia*) of two main traits, sociability and impulsiveness; we can also say that what is demonstrated is the existence of two traits, S and Imp., which are correlated to the extent of 0.5 or thereabouts, thus giving rise to a second-order factor or 'super-factor' of extraversion. The former way of expressing the situation follows the Spearman-Burt-Holzinger line; the latter the Thurstone-Cattell-Guilford line. Logically and statistically these two ways of expressing the existing relationships are, of course, equivalent, and it does not seem necessary to argue in favour of one or the other (Eysenck, 1953). From the practical point of view it would seem that the former solution is preferable, as it requires only two figures (or scores) to denote any individual's position in multifactorial space: his score on E and his score on S - Imp. (or Imp. - S, whichever is preferred). This situation is shown in Fig. 1. The alternative solution requires three scores: S, Imp and E. What cannot be maintained is 'that at least *two independent* (our italics) factors are required to account for the intercorrelations between the E-I variables' (cf. quotation given in the first paragraph of this paper).

The question may be raised why our results and those of Mann (1958) are in some disagreement. The main reason, it may be surmised, is probably the fact that we analysed correlations between single items, while he was concerned with correlations between questionnaires, each of which was made up of many items. Comrey (cf. discussion of his work in Eysenck, 1960c) has shown how untenable the assumption of the univocal nature of such questionnaires can be, in connection

with the MMPI scales, and Eysenck (1956b) has demonstrated that the Guilford S scale, which was constructed to measure a single primary trait, actually split into two independent parts—*introvertive* social shyness and *neurotic* social shyness. Under these conditions it is highly hazardous to use scales of this type for fundamental analyses; the total Guilford S scale correlates with introversion and with neuroticism, but individual items correlate *either* with introversion *or* with neuroticism, but practically never with both. The properties of the scale are therefore quite different from those of its constituent parts, and it is the latter which should be accorded greater prominence in interpretation.

Do our results support the interpretation of E as a unitary dimension? Carrigan does not adequately define the term 'unitary', and its use in factorial theory and practice is none too clear. Guilford would probably claim the status of a unitary trait for his 'S' factor; yet as pointed out above this factor splits up into orthogonal components. Altogether 'primary' factors are often thought of as in some sense more 'unitary' than second-order ones; we shall return to this point in a moment. If by unitary is meant simply 'composed of non-independent constituent units', then our results suggest that E is indeed a unitary factor. If what is meant is 'composed of units related in such a way that their correlations form a matrix of rank one', then clearly E is not a unitary factor. In the second use of the term a unitary factor would correspond more or less with a (Guttman) scaleable universe of items; we know of no non-trivial concept in psychology which meets these requirements. If we accept the criteria suggested by Carrigan herself (p. 330) then it would appear that the evidence from these two studies in favour of the unitary nature of E is favourable on the whole.

As regards the independence of E from the adjustment factor, Carrigan suggests two criteria in her paper: 'factors corresponding to the two dimensions should be uncorrelated, and to the extent that the same variables appear on factors of extraversion-introversion and adjustment, indicators of "good" and "poor" adjustment should as frequently be associated with extraversion as with introversion' (p. 331). The first of these criteria is reasonably well met by our factors, N appearing orthogonal to both E and the S/Imp. factor. (Note also the lack of correlation in our second analysis between scores on N and E.) As regards the second, a search of items having saturations of at least ± 0.25 on E and N revealed four items high on E and N (6, 9, 14, 40); one item high on Introversion and N (33); and one item high on Introversion and high on $(-N)$, i.e. Stability (22). These results are not in line with Carrigan's statement that 'in many instances . . . the characteristics associated with "introversion" continue to have a strong maladjustive flavor' (p. 330). Possibly national differences are in part responsible for differences in 'flavor'. This suggestion is reinforced by the different correlations between S and Imp. on the one hand, and N on the other.

Our findings should not, however, be accepted as indicating any contrary association between E and N. The presence of more items having high loadings on both E and N is, in part, the outcome of a conscious search for such items; in our original work there had been an abundance of items high on Imp. and N, and this resulted in a slight but replicable correlation between these two factors when measured by

the MPI (Eysenck, 1959). In order to counterbalance these items with loadings on both Imp. and N, we collected items which on theoretical grounds were expected to have loadings on E and N. It is now clear that by suitable choice of items complete independence of N and I can be produced. Taking together our various studies in this field (Eysenck, 1956b; Eysenck & Eysenck, 1962) we seem to find that items relating to rigidity are to be found in the dysthymic quadrant (high I, high N) and items relating to aggressiveness in the psychopathic quadrant (high E, high N). Exaggerated impulsiveness and craving for excitement are also found in conjunction with the latter group of items, and 'manifest anxiety' and melancholia in conjunction with the former. It would appear that items can be written to lie at any point in the plane generated by the two factors N and E; this makes the use of 'simple structure' rather less useful in this field than in the cognitive test domain (Eysenck, 1950, 1953). This fact would also seem to explain the repeated slight association on the MPI between Introversion and Neuroticism (Eysenck, 1959); there are more S items in this questionnaire than Imp. items. It may be deduced from our results that had this predominance been reversed, the correlation would have changed sign. By equating the number of items in the two groups perfect orthogonality may be obtained.

Our results may be said to have emphasized the fact that certain subjective judgments play an important (and, we would be prepared to argue, an inevitable) part in the definition and precise location of personality factors in n -dimensional space. Carrigan's two major questions may be thought to carry some scientifically erroneous implications; in asking whether 'extraversion-introversion' is a unitary dimension, and whether 'it' is independent of adjustment, she would seem to be implying the 'real' nature of something called extraversion-introversion. Modern science (and modern philosophy of science) does not work with 'real' entities of this kind in its theoretical analyses. We would prefer to reword the questions to read (albeit somewhat more clumsily): Is it possible to find a combination of questionnaire items such that the answers given to them by a random sample of the population correlate together in a positive manner which can be predicted, roughly at least, on the basis of a psychological theory elaborated by Galen, Kant, Wundt, Jung and other authors (as described by Eysenck, 1960c)? And could the 'factor' thus created be regarded as independent of another combination of questionnaire items answered by the same subjects, items chosen on the basis of clinical and psychiatric experience as representative of 'adjustment'? It seems fairly clear that the possibility of defining two such independent factors does exist, and it remains to explore the utility of these heuristic constructs. Factor analysis cannot, in our submission, prove the 'reality' of these concepts, or produce evidence stronger than permissive; it lacks the *causal* nexus which is essential for reducing the subjective element in the analysis (Eysenck, 1950; 1953, 1956c).

One attempt to provide such a causal theory, and to link it up with experimental laboratory investigations of conditioning, vigilance, reminiscence, figural after-effects, masking, flicker fusion and many other phenomena, has been made by Eysenck (1957, 1960b); the same writer has attempted to link up this theory with a rational system of diagnosis and treatment of neurotic disorders (Eysenck, 1960d, e);

and with the systematic analysis of drug effects (Eysenck, 1962). The success or failure of such attempts will in our opinion be the main determinant of the acceptability of the descriptive terms 'extraversion' and 'neuroticism'; by remaining within the closed circle of factor analysis, writers like Cattell (1957), Cattell & Scheier (1961) and Guilford & Zimmerman (1949), make decisions between their respective systems a matter of subjective judgment, rather than of experimental proof. It is interesting to note that although both these groups of factor-analysts started out with roughly similar pools of items, they have ended up with sets of fifteen or so primary factors which bear no recognizable relation to each other. It is also notable, however, that second-order factor analyses in each case have given rise to strongly-marked factors of extraversion-introversion and neuroticism (called 'anxiety' by Cattell & Scheier (1961)), thus suggesting that these higher-order factors emerge as independent entities almost regardless of the course of analysis followed, the pool of items used originally, or the populations of subjects studied. (Cf. Bargatta's, 1962, paper in this connection.)

An explanation of these findings can be made in terms of Eysenck's original theory (1957). According to this theory, the 'traits' which constitute the primary factors of Guilford and Cattell alike are essentially habits (SH_R), and habits are notoriously shifting, difficult to classify and pin down; indeed, any such classification would have little objective basis. Extraversion-introversion and neuroticism, however, are conceived in much more fundamental terms as constitutionally-determined properties of the C.N.S. and the autonomic system, respectively, determining such vitally important reaction-properties of the organism as speed of conditioning, strength of autonomic reaction, etc.; small wonder that they emerge from any full-scale attempt to analyse personality.

It is not suggested, of course, that the constitutional factors underlying extraversion and neuroticism are not modified by experience and learning; this point has been discussed in considerable detail by Eysenck (1960a). Experimental work by Cattell *et al.* (1955, 1957) and Eysenck (1956) and Eysenck & Prell (1951) suggests that both heredity and environment do play an important part in these broad behaviour patterns, and even such apparent differences in emphasis as Cattell's stress on *environment* in the causation of extraverted behaviour, as opposed to Eysenck's stress on *heredity*, may be reconciled by reference to the dual nature of this concept; it is not inconceivable that sociability is more easily subject to environmental control, while impulsiveness may have deeper roots in heredity. This, of course, is mere speculation; it may, however, suggest a way of reconciling apparently contradictory results, and also emphasize the importance of recognizing the dual nature of extraversion. In all this there is, of course, no implication that extraversion is composed of no more traits than the two investigated here; this is an empirical problem, the answer to which is not to be anticipated.

REFERENCES

BARGATTA, E. F. (1962). The coincidence of subtests in four personality inventories. *J. soc. Psychol.*, **56**, 227-244.

CARRIGAN, P. M. (1960). Extraversion-introversion as a dimension of personality: a reappraisal. *Psychol. Bull.* **57**, 329-360.

CATTELL, R. B. (1957). *Personality and Motivation Structure and Measurement*. New York: World Books.

CATTELL, R. B., BLEWETT, P. B. & BELOFF, J. R. (1955). The inheritance of personality: a multiple variance analysis determination of approximate nature-nurture ratios for primary personality factors in Q-data. *Amer. J. hum. Genet.*, 7, 122-146.

CATTELL, R. B., STICE, G. F. & KINESTY, N. F. (1957). A first approximation to nature-nurture ratios for eleven primary personality factors in objective tests. *J. abnorm. soc. Psychol.*, 54, 143-159.

CATTELL, R. B. & SCHEIER, I. H. (1961). *The Meaning and Measurement of Neuroticism and Anxiety*. New York: Ronald Press.

COHEN, J. (1959). *Chance, Skill and Luck*. Harmondsworth: Penguin Books Ltd.

EYSENCK, H. J. (1950). Criterion analysis—an application of the hypothetico-deductive method to factor analysis. *Psychol. Rev.*, 57, 38-53.

EYSENCK, H. J. (1953). The logical basis of factor analysis. *Amer. Psychologist*, 8, 105-114.

EYSENCK, H. J. (1956a). The inheritance of extraversion-introversion. *Acta psychol.*, 12, 95-100.

EYSENCK, H. J. (1956b). The questionnaire measurement of neuroticism and extraversion. *Riv. Psicol.*, 50, 113-140.

EYSENCK, H. J. (1956). L'analyse factorielle et le problème de la validité. In: *L'analyse factorielle et ses applications*. Paris: Centre National de la Recherche Scientifique.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. New York: Praeger.

EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: Univ. of London Press.

EYSENCK, H. J. (1960a). Levels of personality, constitutional factors and social influences: an experimental approach. *Internat. J. soc. Psychiat.*, 6, 12-24.

EYSENCK, H. J. (ed.) (1960b). *Experiments in Personality*. New York: Praeger.

EYSENCK, H. J. (1960c). *The Structure of Human Personality* (2nd Ed.). New York: Macmillan.

EYSENCK, H. J. (Ed.) (1960d). *Behaviour Therapy and the Neuroses*. New York: Pergamon Press.

EYSENCK, H. J. (1960e). A rational system of diagnosis and therapy in mental disease. In: *Progress in Clinical Psychology* (eds. L. E. Abt & R. Riess), IV, 46-64. New York: Grune & Stratton.

EYSENCK, H. J. (Ed.) (1962). *Experiments with Drugs*. New York: Pergamon Press.

EYSENCK, H. J. & Claridge, G. (1962). The position of hysterics and dysthymics in a two-dimensional framework of personality description. *J. abnorm. soc. Psychol.*, 64, 46-55.

EYSENCK, H. J. & Prell, D. B. (1951). The inheritance of neuroticism: an experimental study. *J. ment. Sci.*, 97, 441-465.

EYSENCK, S. B. G. (1962). The validity of a personality questionnaire as determined by the method of nominated groups. *Life Sciences*, 1, 13-18.

EYSENCK, S. B. G. & EYSENCK, H. J. (1962). A factorial study of an interview-questionnaire. *J. Clin. Psychol.*, 18, 286-290.

EYSENCK, S. B. G. & EYSENCK, H. J. (1962). Rigidity as a function of introversion and neuroticism. *Internat. J. soc. Psychiat.*, 8, 180-184.

EYSENCK, S. B. G. & EYSENCK, H. J. (1962). The validity of questionnaires and assessments of extraversion and neuroticism and their factorial stability. In press.

GUILDFORD, J. P. & GUILFORD, R. B. (1934). An analysis of the factors in a typical test of introversion-extraversion. *J. abnorm. soc. Psychol.*, 28, 377-399.

GUILFORD, J. P. & ZIMMERMAN, W. S. (1949). *The Guilford-Zimmerman Temperament Test*. Beverley Hills, Calif.: Sheridan Supply Co.

MANN, R. D. (1958). The relationship between personality characteristics and individual performance in small groups. Unpublished doctoral diss., Univ. of Michigan. (Quoted by Carrigan, 1960.)

SINGER, J. L., WILENSKY, H. & McCRAVEN, V. G. (1956). Delaying capacity, fantasy, and planning ability: a factorial study of some basic functions. *J. cons. Psychol.*, 20, 375-383.

A Study of Neuroticism and Casual Arterial Blood Pressure*

BY J. O. ROBINSON

*M.R.C. Social Psychiatry Research Unit, Detachment at Llandough Hospital,
Penarth, Glamorgan†*

Tests of neuroticism were given to a random sample of 167 individuals from a geographically defined population, whose blood pressures had been measured under standard conditions. No significant correlation was found between blood pressure and a combined test score for neuroticism except in the 41-50 age group. In this group there was a significant correlation between systolic blood pressure and the combined test score, but the numbers were small and the correlation could have been fortuitous.

A group of 56 out-patient neurotics made significantly higher scores on the combine measure than did the random sample. A group of 43 out-patients with diastolic blood pressures of more than 100 mm Hg (in whom secondary hypertension was not suspected) also made significantly higher scores than the random sample. Since, in most age groups, the correlation between blood pressure and neuroticism was low, it is possible that the data could have resulted from selection of patients, probably by the general practitioner. Such selection could have had the effect of presenting to the hospital only those hypertensives who also had high neuroticism scores.

I. INTRODUCTION

The common clinical impression of the coincidence of hypertension with certain personality characteristics has received attention both from experimenters and clinicians. Whilst positive claims have been made serious criticism can be levelled at the methods used and the conclusions drawn.

It is of little value to describe a personality type and associate it with hypertension unless it can be shown that a significantly higher proportion of hypertensives than normotensives belong to the type.

Gressel, Shobe, Saslow, DuBois & Schroeder (1949) pointed out the frequent lack of normotensive controls in work on personality in hypertension. In fact, very few studies have been done completely without controls (Binger, Ackerman, Cohn, Schroeder & Steel, 1945; Hambling, 1951; Wolf, Cardon, Shepard & Wolff, 1955) but many have used controls in unusual ways. Palmer (1950) mentioned controls but it is difficult to see whether or not his descriptions of the personalities of hypertensives were derived by means of comparison with his controls. The same can be said of Wolf & Wolff (1950) and many other descriptions of personality in hypertensives.

Controls are useful only if the method of personality assessment is objective and standard or, at least, 'blind' in the sense that the assessors know nothing about the blood pressures of the subjects. An objective method is preferable if observations are to be repeatable by other workers.

* Based on a thesis approved for the degree of Ph.D. in the University of London (1961).

† Part of this work was carried out whilst the author was a member of the M.R.C. Neuropsychiatric Research Unit.

Gressel *et al.* (1949) rated subjects on six characteristics; impulsiveness, subnormal assertiveness, obsessive-compulsive behaviour, depressive behaviour, anxiety and hysteria. Five degrees of severity were defined for each characteristic to make the assessment as objective as possible. However, the data for the rating were obtained by psychiatric interview. If different interviewers had done the work, different information might have been obtained. If the assessments had been blind then this objection would lose its force, but they were not. Many of the hypertensives were rated before work was begun on the two other groups (personality disorders and somatic disorders). Thus, in many cases the interviewer knew that the subject was hypertensive and any presuppositions he might have had could have biased his assessment.

The hypertensive group resembled the group with personality disorders on the ratings of all but obsessive-compulsive behaviour and subnormal assertiveness. This finding has not been independently confirmed and the method used would make such confirmation difficult.

If personality is not assessed objectively, a blind assessment may avoid observer bias. But, whilst a difference in personality between hypertensive and normotensive subjects may be demonstrated, a reliable description of personality is not necessarily arrived at by such means. A series of papers by Harris, Sokolow and their associates (Harris, Sokolow, Carpenter, Freedman & Hunt, 1953; Kalis, Harris, Sokolow & Carpenter, 1957; Kalis, Harris, Bennett & Sokolow, 1961) illustrates this. These workers made use of a list of 389 adjectives. Those which were considered by observers to apply to the behaviour of the subject during psychodrama were ticked.

If observers are naive and there is no obvious difference between the experimental groups, then an unbiassed personality description should be obtained by this method. If, however, there is some obvious general difference between control and experimental groups, as was the case in these studies (this was shown by the success of the 'overall prediction' based on interview), and the observers expect the hypertensive to show certain personality characteristics, then they may tend to tick adjectives according to their own presuppositions. If they all expect to find the same personality characteristics in hypertensives, then the results of the check lists will be consistent, but spuriously so. A difference between the experimental and control groups will have been demonstrated but the nature of the difference will remain obscure.

Sainsbury (1960) used the Maudsley Personality Inventory for the assessment of personality and thus avoided difficulties of description and definition. The neuroticism scale of this inventory has been shown many times (Eysenck, 1959) to distinguish between neurotic and non-neurotic groups. In Sainsbury's study hypertensive out-patients showed significantly higher neuroticism scores than control patients.

These studies have claimed to demonstrate typical personality characteristics in individuals with high blood pressure. Other published work has dealt with elevations of blood pressure during emotional stress induced during an interview (for example Hambling, 1958; Wolf, Pfeiffer, Ripley, Winter & Wolff, 1948; Van der Valk, 1957), or by the performance of various tasks (for example Brod, 1959; Innes,

Millar & Valentine, 1959; Imhof, 1957). Other investigators have applied stressful stimuli, for example Jost, Ruilmann, Hill & Gulo (1952) who asked questions 'which may have had emotional connotation for the subject', and Schachter (1957) who created frightening and annoying situations. In these investigations elevations of blood pressure took place in both normotensives and hypertensives but the rise in hypertensives was much greater, both absolutely and as a proportion of the original level.

Various conclusions were drawn about the emotional characteristics of hypertensives, but the differences in elevation of blood pressure cannot be regarded as unequivocal evidence that emotional differences exist, for the hypertensive also reacts to physical stimuli and drugs in an exaggerated way. In any case the regulation of blood pressure is such a complex process that elevation during some kind of stimulation might be almost any function of the original level. It need not be a simple one. The increment may be greater as a proportion of the original level in those individuals with higher original levels and this may be quite independent of emotional characteristics.

Nevertheless, it seems likely that emotional difficulty may be one factor in the pathogenesis of hypertension. There is evidence that trauma (Ruskin, Beard & Schaffer, 1948), battle experience (Graham, 1958) and change from 'primitive' to 'civilized' societies (Cruz-Coke, 1960) can give rise to lasting elevations of blood pressure. There are also indications that animals are differentially susceptible to hypertension induced by noxious stimulation. In some cases this seems to be a function of breed (Yeakel, Shenkin, Rothballe & McCann, 1948).

2. THE PRESENT STUDY

The present work is part of a longitudinal study of blood pressure in a random sample from the complete population of the Rhondda Fach, a mining valley in South Wales (Miall & Oldham, 1955, 1958; Miall, 1959). It was decided to investigate relationships between neuroticism and blood pressure in this sample and to compare neuroticism scores with those for hypertensives and neurotics.

Neuroticism was chosen for a number of reasons.

1. Several writers have pointed out the similarity of symptoms of neurosis and uncomplicated hypertension (Ayman & Pratt, 1931; Riseman & Weiss, 1930; Hopkins, 1958).
2. The tests could be independently validated. A group of hospital out-patient neurotics was included for this purpose.
3. The tests available were objective, simple, easily transportable and easily administered in the subjects' homes.

One of the tests (the Maudsley Personality Inventory) included an extraversion scale which had been shown (Eysenck, 1959) to distinguish significantly between groups of hysterics and dysthymics. This, and the Porteus Maze Test, which also has distinguished these two groups (Foulds, 1951), were included. No significant relationships were found between blood pressure and these scores. The scores of the hypertensives were not distinguishable from those of the random group.

3. THE SUBJECTS

1. The Random Group

A group had been selected in 1954 (see Miall & Oldham, 1955) for the first survey of the longitudinal study of blood pressure. A private census of the population of the Rhondda Fach had been recorded on cards which were numbered serially beginning with the head of the household in the first street (alphabetically) of the first town and ending with the youngest member of the household of the last street of the last town. From these records a group of 261 people had been selected by extracting every card bearing a number ending with 77. Of this sample, 250 agreed to take part in the first survey. Table 1 shows the position at the time of the second survey when the psychological testing was done. For the psychological tests age limits of 13 and 65 were applied; 175 subjects between these limits were still resident in the valley.

Table 1. *Subjects*

	Investigated	Full sets of data	Too ill	Refused	Total
Random group	167	162	2	6	175
Neurotic group	56	53	0	4	60
Hypertensive group	43	33	3	3	49

(For a number of reasons full sets of data were not obtained for all subjects. These were omitted from the analysis.)

2. The Neurotic Group

From the psychiatric clinic of a hospital serving the area, a number of neurotic patients were selected who fulfilled the following criteria:

- (a) No psychotic or organic mental illness was suspected.
- (b) The patients were out-patients living in the Rhondda or neighbouring valleys.
- (c) They were within the age limits.

Patients' blood pressures were not known at this stage.

3. The Hypertensive Group

The records of all patients attending two medical clinics at the same hospital over a period of 6 months were searched and a group was selected according to the following criteria:

- (a) Their diastolic blood pressures as measured at the clinic were equal to or exceeded 100 mm.
- (b) Their hypertension was not known to be secondary to another disorder.
- (c) They were out-patients living in the Rhondda or neighbouring valleys.
- (d) They were within the age limits.

No deliberate elimination of neurotics took place.

4. PROCEDURE

(a) Blood pressure was measured in the subjects' homes. Miall did the measurements for the random sample in the way described by Miall & Oldham (1955). The writer did the measurements for the neurotic group after trials had established that his measurements corresponded closely to Miall's for the same subjects (correlations were about 0.9 for both systolic and diastolic pressures and differences between means were less than 4 mm Hg). For the hypertensive group, the values appearing in the clinic records were used.

(b) The psychological tests were conducted at the subjects' homes by the writer. The tests were chiefly selected from those described by Eysenck, Granger & Brengelmann (1957). The 'situations questions' consisted of a list of eight questions about how subjects felt in certain situations which might be expected to be stressful. A direct 'Yes' or 'No' in the scoring direction, that is in the direction which indicated that the subject would find the situation stressful, was scored. Anything else was ignored. As far as possible the tests were

arranged so that pencil and paper tests alternated with others. Thus the tests, in order of presentation, were:

1. Maudsley Personality Inventory (two scores; one for neuroticism and one for extraversion).
2. U.S.E.S. Manual Dexterity Test.
3. Pressey Cross-out List of Annoyances.
4. Pressey Cross-out List of Interests and Dislikes (two scores, one for interests and one for dislikes).
5. Measure of Static Ataxia.
6. List of Food Dislikes.
7. Porteus Mazes.
8. Pressey Cross-out List of Worries.
9. Situations Questions.

5. RESULTS

A discriminant function (Fisher, 1936) was calculated using nine of the test scores (those which were expected to measure neuroticism) of those subjects in the random and neurotic groups who were over 40 years of age. The reason for the age restriction was the paucity of data for hypertensives under 40 years of age. Since a different discriminant function might have been obtained had the younger age group been included, it was considered safer to confine this part of the analysis to the older subjects.

'Neuroticism scores' were defined as the combination obtained by adding the test scores weighted by means of the discriminant function. The mean values for the neuroticism scores for subjects over 40, and the levels of probability associated with the differences between groups, are shown in Table 2.

Table 2. *Differences between neuroticism* scores for the subjects over 40 years of age*

	Group size	Mann-Whitney	
		<i>p</i>	<i>t</i>
Random/Neurotic	60/25	0.00006	5.33 <i>p</i> < 0.001
Random/Hypertensive	60/31	0.0068	3.60 <i>p</i> < 0.001
Hypertensive/Neurotic	31/25	0.1116	1.48 N.S.

	Means
Random	51.37
Neurotic	70.64
Hypertensive	63.91

* The weighted combination of test scores.

The neuroticism score of the hypertensive group was not significantly different from that of the neurotic group and was significantly greater than that of the random group.

Next, correlations were calculated between neuroticism score and blood pressure for the whole of the random sample. Blood pressures were corrected for age using the method described by Miall & Oldham (1955). The adjusted value obtained by this method represents the deviation of an individual's pressure from the mean for his age group, further adjusted to equalize the standard deviations in each age and sex group.

Table 3. *Correlations with neuroticism**

Age	Sex	No.	Systolic blood pressure	Diastolic blood pressure	Age-adjusted systolic	Age-adjusted diastolic	Age
41-50	M	14	+0.59†	+0.34	+0.63‡	+0.29	+0.35
	F	15	+0.33	+0.36	+0.31	+0.33	+0.13

* The weighted combination of test scores.

† $p < 0.05$.

‡ $p < 0.01$.

The only significant correlation coefficients occurred in the male 41-50 age group where the values for systolic blood pressure were significant (see Table 3). Thus any group of males with high systolic blood pressures selected from this age group would be expected also to show high neuroticism scores.

6. DISCUSSION

The high positive correlations in the 41-50 age group are interesting findings, but the number of subjects involved is small compared with the total studied and the number of tests of significance was large. It would not seem unreasonable to dismiss these findings as due to chance. If the correlation is not confirmed¹ by future work then the large difference between the neuroticism scores of the random and hypertensive groups cannot be explained purely in terms of blood pressure. Indeed, it is possible by examining the present results in more detail to see that such an explanation may not be very satisfactory even here.

If the high scores of the hypertensive group of the present work were accounted for by blood pressure alone, then correction of neuroticism scores for blood pressure should eliminate the difference in score between the random and hypertensive groups. Such a correction has been applied (Table 4) and does not eliminate

Table 4. *Correction of neuroticism* scores for blood pressure level*

Group size	Random group (over 40)		Hypertensive group (over 40)	
	M	F	M	F
Raw mean neuroticism	50.37	30	15	16
Mean systolic blood pressure	144.33	52.33	62.73	64.88
Adjusted mean neuroticism	52.57	53.78	60.13	60.47
Covariance 'F' value			2.80 N.S.	
Mean diastolic blood pressure	88.33	87.00	109.00	108.13
Adjusted mean neuroticism	52.15	54.48	58.81	61.20
Covariance 'F' value			2.26 N.S.	

* The weighted mean of the test scores.

the difference, though the corrected values are not significantly different from each other. Bearing in mind that the measurements of blood pressure are not strictly other.

¹ Since the time of writing a further group of 148 subjects in the 41-50 age group has been tested. The correlation between neuroticism score and blood pressure has not been confirmed.

comparable and that those made on the hypertensives in the unusual surroundings of the clinics are probably rather high (Miall & Oldham, 1958), these scores may be overcorrected. Thus the hypertensive group may be more neurotic than would be expected from its blood pressure distribution. Moreover a subgroup of members of the random group aged over 40 can be selected whose diastolic blood pressures were 100 mm or more (comparable, that is, with those of the hypertensive group). This subgroup had a mean neuroticism score of 53.79 as against 63.91 for the hypertensive group.

Two further subgroups can be chosen so that the systolic blood pressures are more closely comparable. The eleven members of the hypertensive group whose age-adjusted diastolic blood pressures are less than +20 have mean age-adjusted pressures of 21.36/6.36 and the ten members of the random group whose age-adjusted diastolic blood pressures are greater than or equal to +20 have mean age-adjusted pressures of 22.50/26.50. The mean neuroticism scores of these two subgroups are 62.73 and 53.68 respectively, even though the mean diastolic blood pressure of the latter is appreciably higher.

The high neuroticism scores of the hypertensives could have been due to the circumstances under which the measurements were taken:

1. The effect on both blood pressure and neuroticism score of attendance at hospital is unknown. All the psychological testing was done at home, but it is possible that the hypertensives, had they been given the psychological tests before they became hospital out-patients, might have shown neuroticism scores more like those of the random group. The worry associated with waiting for and making a visit to the hospital could have influenced the scores.

2. The hypertensive's knowledge of his own hypertension may also be important. Stewart's work (1953) suggested that the patient's awareness of his hypertension greatly influenced the occurrence of headache, for instance. It is not known which of the hypertensives were aware of their high blood pressure.

3. Many of the hypertensives had been treated with rauwolfa derivatives; such drugs are considered to have psychological side-effects and it is not known whether the results of the tests in the present work were affected.

It is possible that much of the high score of the hypertensive group was attributable to its neurotic members who need not have been sustained hypertensives. They could have been subjects who gave a high reading only in the clinic situation. Eysenck (1960), in a review of work dealing generally with autonomic lability, identified the factor of lability found by various workers with his factor of neuroticism. Other writers (Ostfeld & Lebovitz, 1959; Innes *et al.*, 1959; Malmo & Shagass, 1952) have presented evidence of lability of blood pressure in neurotic groups. In the present work the blood pressure value probably represented an average value raised by some amount characteristic of the individual. In a person with a very stable blood pressure the measurement would probably be a good estimate of average blood pressure. On the other hand, in a person with labile blood pressure it would not. Thus the subjects in the present work whose blood pressure values were high were probably both those with high average blood pressures and those with labile blood pressures. Further work will be required to

separate these two groups, but the inclusion of labile individuals in the hypertensive group would seem to be likely to swell its neuroticism score.

Many of the symptoms of hypertension are similar to those of emotional disorders, as has been pointed out by several writers mentioned earlier. Of the patients who come to the hospital with the symptoms common to hypertension and neurosis, those with high blood pressure may be directed to the medical clinic and those without such a definite sign will probably go to the psychiatric clinic.

Thus the observations could be explained in terms of selection either by the general practitioner, or by the patient himself in his decision to consult a doctor about a particular symptom, without the necessity of any causal relationship between neuroticism and hypertension.

The writer is indebted to Professor A. L. Cochrane for his help at all stages of the work, to Dr W. E. Miall for allowing his blood pressure measurements to be used, to Dr I. T. T. Higgins for help in selecting the hypertensives and to Mr P. D. Oldham for statistical advice. The co-operation of Dr M. W. Annear, Dr T. S. Davies, Dr E. H. Evans, the Secretary, doctors and members of the records staff of the East Glamorgan Hospital and of the people of the Welsh valleys is also gratefully acknowledged.

Computation was carried out on the 'Zebra' computer of Standard Telephones and Cables, Ltd., to whom the writer's thanks are also due.

REFERENCES

AYMAN, D. & PRATT, J. H. (1931). Nature of the symptoms associated with essential hypertension. *Arch. intern. Med.* **47**, 675-687.

BINGER, C. A. L., ACKERMAN, N. W., COHN, A. E., SCHROEDER, H. A. & STEEL, J. M. (1945). *Personality in arterial hypertension*. Psychosomatic Medicine Monographs. New York: American Society for Research in Psychosomatic Problems.

BROD, J., FENCL, V., HEJL, Z. & JIRKA, J. (1959). Circulatory changes underlying blood pressure elevation during acute emotional stress (mental arithmetic) in normotensive and hypertensive subjects. *Clin. Sci.* **18**, 269-279.

CRUZ-COKE, R. (1960). Environmental influences and arterial blood pressure. *Lancet*, **2**, 885-886.

EYSENCK, H. J., GRANGER, G. W. & BRENGELMANN, J. C. (1957). *Perceptual processes and mental illness*. London: Chapman and Hall, Ltd.

EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory*. London: University of London Press.

EYSENCK, H. J. (1960). *The structure of human personality* (2nd edition). London: Methuen.

FISHER, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugen., Lond.* **7**, 179-188.

FOULDS, G. A. (1951). Temperamental differences in maze performance. *Brit. J. Psychol.* **42**, 209-217.

GRESSEL, G. C., SHOBE, F. O., SASLOW, G., DUBoIS, P. H. & SCHROEDER, H. A. (1949). Personality factors in arterial hypertension. *J. Amer. med. Ass.* **140**, 265-275.

HAMBLING, J. (1951). Emotions and symptoms in essential hypertension. *Brit. J. med. Psychol.* **24**, 242-253.

HAMBLING, J. (1958). Essential hypertension. In *The Nature of Stress Disorder*. Conference of the Society for Psychosomatic Research. London: Hutchinson Medical Publications.

HARRIS, R. E., SOKOLOW, M., CARPENTER, L. G., FREEDMAN, M. & HUNT, S. P. (1955). Response to psychologic stress in persons who are potentially hypertensive. *Circulation*, **7**, 874-879.

HOPKINS, P. (1958). Hypertension in general practice. *Practitioner*, **180**, 463-472.

IMHOF, P., HURLIMANN, A. & STEINMANN, B. (1957). Über blutdrucksteigerung bei psychischer belastung. *Cardiologia*, **32**, 272-276.

INNES, G., MILLER, W. M. & VALENTINE, M. (1959). Emotion and blood pressure. *J. ment. Sci.* **105**, 840-851.

JOST, H., RUILMANN, C. J., HILL, T. S. & GULO, M. J. (1952). Studies in hypertension; central and autonomic nervous system reactions of hypertensive individuals to simple physical and psychologic stress situations. *J. nerv. ment. Dis.* **115**, 152-162.

KALIS, B. L., HARRIS, R. E., SOKOLOW, M. & CARPENTER, L. G. (1957). Response to psychological stress in patients with essential hypertension. *Amer. Heart J.* **53**, 572-578.

KALIS, B. L., HARRIS, R. E., BENNETT, L. F. & SOKOLOW, M. (1961). Personality and life history factors in persons who are potentially hypertensive. *J. nerv. ment. Dis.* **132**, 457-468.

MALMO, R. B. & SHAGASS, C. (1952). Studies of blood pressure in psychiatric patients under stress. *Psychosom. Med.* **14**, 82-93.

MIALL, W. E. & OLDHAM, P. D. (1955). A study of arterial blood pressure and its inheritance in a sample of the general population. *Clin. Sci.* **14**, 459-488.

MIALL, W. E. & OLDHAM, P. D. (1958). Factors influencing arterial blood pressure in the general population. *Clin. Sci.* **17**, 409-444.

MIALL, W. E. (1959). Follow up study of arterial pressure in the population of a Welsh mining valley. *Brit. med. J.* **2**, 1204-1210.

PALMER, R. S. (1950). Psyche and blood pressure; 100 mental stress tests and 50 personality surveys in patients with essential hypertension. *J. Amer. med. Ass.* **144**, 295-298.

RISEMAN, J. E. F. & WEISS, S. (1930). The symptomatology of arterial hypertension. *Amer. J. med. Sci.* **180**, 47-59.

RUSKIN, A., BEARD, O. W. & SCHAFER, R. L. (1948). 'Blast hypertension' Elevated arterial pressures in victims of the Texas City disaster. *Amer. J. Med.* **4**, 228-236.

SCHACHTER, H. (1957). Pain, fear, and anger in hypertensives and normotensives; a psychophysiological study. *Psychosom. Med.* **19**, 17-29.

STEWART, I. McD. G. (1953). Headache and hypertension. *Lancet*, **1**, 1261-1266.

VAN DER VALK, J. M. (1957). Blood-pressure changes under emotional influences, in patients with essential hypertension, and control subjects. *J. psychosom. Res.* **2**, 134-146.

WOLF, S., PFEIFFER, J. B., RIPLEY, H. S., WINTER, O. S. & WOLFF, H. G. (1948). Hypertension as a reaction pattern to stress; summary of experimental data on variations in blood pressure and renal blood flow. *Ann. intern. Med.* **29**, 1056-1076.

WOLF, S. & WOLFF, H. G. (1951). Life stress and essential hypertension. In, E. T. Bell ed., *Hypertension*. Minneapolis: University of Minnesota Press.

WOLF, S., CARDON, P. V., SHEPARD, E. M. & WOLFF, H. G. (1955). *Life stress and essential hypertension: a study of circulatory adjustments in man*. Baltimore: The Williams and Wilkins Co.

YEAKEL, E. H., SHENKIN, H. A., ROTHBALLER, A. B. & McCANN, S. McD. (1948). Blood pressure of rats subjected to auditory stimulation. *Amer. J. Physiol.* **155**, 118-127.

Manuscript received 20 October 1961

Revised manuscript received 7 February 1962

A Method for Testing 'Simulated' Visual Field Defects

By JAMES INGLIS AND H. A. MACKAY
Queen's University, Kingston, Ontario

It was required to discover if a patient's complaint of blindness in the upper field of vision were simulated or not. The results of an attempt to investigate this problem by conditioning techniques proved ambiguous. A technique was devised which gave unequivocal results.

This brief report is concerned with the investigation of a complaint of a visual field defect in a 59-year-old man. This patient complained of being totally blind in the whole upper half of his visual field. When he was seen by the psychologist he had already undergone standard perimetric investigation and it had been shown that he failed to report on the position of the stimulus light whenever this fell anywhere in the upper quadrants. Since no definite organic lesion could be shown to account for this he was referred for investigation of the possibility that he was malingering and that this partial blindness was, in fact, only simulated.

The psychologist first of all confirmed the perimetric findings. He found consistent impairment of vision for targets of three sizes, three colours and three brightness values. Further confirmation of the patient's failure to report the appearance of stimuli projected on some visual areas was obtained by means of the Harrington-Flocks Multiple Pattern Visual Field Screener (Harrington & Flocks, 1954).

Neither of these techniques can, however, differentiate between inability to perceive and failure to report perception. Means had to be devised, therefore, to separate these two aspects of the patient's performance.

The first method employed a conditioned psychogalvanic response. Electroshock provided the unconditioned stimulus for this response and the perimeter light was used as the conditioned stimulus. It was expected that if the patient's disorder were due merely to a failure accurately to report visual perception then a psychogalvanic response should inadvertently appear even when the stimulus light was projected on the ostensibly 'blind' upper half of his visual field, after a number of reinforced conditioning trials had been given. No such response should be elicited if there were a real failure to perceive.

Two conditioning sessions were conducted. The first session included ten unreinforced test trials in which the conditioned stimulus alone was presented. On half of these test trials the stimulus was projected upon the lower part of the visual field and in the remaining trials it was projected on the upper ('blind') part. In this session a conditioned response was successfully obtained on five out of the five test trials with the stimulus in the lower half of the patient's visual field. A positive response was, on the other hand, obtained on only two of the five occasions in which the stimulus fell in the upper quadrants. This result might seem to support the notion that the patient actually failed to perceive the stimulus in the

latter trials but it was, of course, not possible to assess the true significance of the difference between these sets of results. A second session was therefore undertaken in order to find out at least how repeatable these results might be. This attempt was, however, invalidated by the difficulty on this occasion of securing any autonomic response at all even to very strong electroshock.

In any case it proved very difficult indeed to ensure that the patient kept a constant fixation during the test trials, which were carried out in a dark room. It was evident that eye movements alone might produce spuriously positive responses. Another mode of assessment therefore had to be devised.

This method required the patient unwittingly to inspect the perimetric target through an inverting lens. Perimetric tests were then again carried out in a darkened room so that the patient was unaware that his retinal image was, in fact, inverted. Since the inspection target provided a field uniform in each quadrant it could not itself give cues to orientation.

It was then expected that if the defect were simulated the patient would still report himself unable to see the stimulus light in the upper quadrants. Because of the inverting lens the light would then, in reality, be projected upon the lower ('intact') half of his visual field. He would also, if malingering, be expected to report the light when it was actually projected on the upper half, when it would appear to him to strike the lower quadrants of his vision.

If, on the other hand, his blindness were *not* simulated the inverting lens would invert the defect, causing him to report stimuli apparently in the formerly blind upper half but not in the formerly intact lower half of his visual field.

On testing it was found that the patient reported an inversion of his defect under these conditions, thus behaving in accordance with the notion that his partial blindness was *not* simulated.

Subsequent search of the ophthalmological literature (Duke-Elder, 1949) showed that similar but apparently not identical tests based on optical principles have previously been employed for the detection of other forms of visual malingering.

REFERENCES

DUKE-ELDER, W. S. (1949). *Text-book of Ophthalmology*. London: Henry Kimpton.
HARRINGTON, D. O. & FLOCKS, M. (1954). Visual field examination by a new tachistoscopic multiple pattern method. *Amer. J. Ophthal.*, 37, 719.

Manuscript received 20 July 1962

A Reply to Dr Emery

By ALAN RICHARDSON

*Department of Psychology,
University of Western Australia*

In reviewing the writer's monograph on 'The Assimilation of British immigrants in a Western Australian Community' in the first issue of this journal Dr Emery appears to have based most of his major criticisms on certain misunderstandings of the theory and methods that have been employed. This reply will confine itself to these major criticisms, as any discussion of the more general issues of assimilation that are raised by Dr Emery would rapidly reach book-size proportions.

In referring to the overall scale of assimilation Dr Emery asserts that, 'neither the facts in this report nor those presented in R. Taft's recent replication (1961) provide encouraging support for the conceptual scheme or for the measuring instruments'.

The first piece of evidence that is presented to support this statement is said to be the failure of the assimilation scale 'to show a significant relation with length of time in Australia or with age of the migrant'. As a clarification of the theory is a necessary background to answering this and some of the other criticisms it will be useful to give a brief restatement of its main features. The basic proposition in this theory states that the process to be called assimilation may be defined as having three interrelated and measurable stages. These three stages consist in a satisfaction stage, an identification stage and an acculturation stage. 'The starting point in this progression depends upon the immigrant reaching a specifiable level of satisfaction with life in Australia. If important needs are not satisfied then the immigrant not only fails to arrive at the point where further assimilation may occur but as a result of not satisfying these needs may retreat to a condition of dissatisfaction with his life in Australia. Given this specifiable level of satisfaction, however, it becomes possible, but not inevitable, that the immigrant will begin to think of himself as to some extent Australian. This stage of identification will vary in its intensity from immigrant to immigrant. Among those who experience a relatively strong identification with Australia there will be some who, in the presence of certain specifiable conditions, are likely to move on to the final stage of assimilation which involves actually becoming like Australians of similar background. This final stage of relatively high level acculturation is not automatically achieved by everyone who reaches the identification stage' (Richardson, 1961, p.29).

From this quotation it should be clear that an immigrant who, for example, arrives in Australia late in life might become satisfied or dissatisfied whether he continued to live in Australia for 5 or 50 more years. To reach the identification stage, however, and even more to reach the stage of acculturation both time and youth are advantageous. For example, among the male immigrants there are only 4 out of the 30 identified migrants who have reached this stage of identification

within their first three years in Australia, while not one of the 19 acculturated migrants has lived in Australia for so short a time and only 5 of the 19 have lived in Australia as little as 5 or 6 years. Nine of the 19 acculturated immigrants have lived in Australia for at least 9 years.

However, Dr Emery seems to feel that because those immigrants who are acculturated have been in Australia for a long time, that all those immigrants who have been in Australia for a long time must be acculturated. The facts, however, do not support his belief, and my conceptual scheme makes no such assumptions.

Dr Emery's second piece of evidence against the conceptual scheme and the assimilation scale comes from the fact that only 17 of the 30 immigrants (men and women combined) who score high on the acculturation subscale are also satisfied and identified. Though this is true, and it would be a much better scale if it were otherwise, it is not entirely relevant to select one particular set of errors for comment without considering the reliability and usefulness of the entire assimilation scale. That the frequency of non-scale patterns (errors) was relatively low on this assimilation scale is shown by the Reproducibility Index of 0.94 for the husbands ($N = 90$) and 0.97 for the wives ($N = 90$) with respective Jackson Plus Percentage Ratios of 0.80 and 0.88. Its usefulness is shown by the many relationships between assimilation level and other variables that were reported in the monograph.

In discussing the concept of satisfaction and the subscale used for measuring it, a second area of confusion seems to be present. Dr Emery is worried by the apparent emphasis on the 'material conditions' of life as determinants of satisfaction and goes on to say that, 'identification may well develop even though a person is dissatisfied with his daily lot; and . . . a high degree of satisfaction with material conditions need not lead to identification'. As we have seen the truth of this latter assertion is not in dispute—satisfaction does not necessarily lead on to identification. However, the first part of the quotation seems to be based on a misunderstanding of the conceptual significance of the term satisfaction. With the present sample of immigrants it is not true that, 'identification may well develop even though a person is dissatisfied with his daily lot', i.e. material conditions. But if hypotheses are being considered as to what makes for satisfaction among refugee immigrants, or among other groups such as voluntary professional people coming to Australia, it may well be that something other than material conditions will play an important part. Perhaps it is the phrase 'typical immigrant' (Richardson, 1961, p. 50), quoted by Dr Emery, that has led him to suppose that a generalization was being made to all immigrants, though the context should have been enough to indicate that reference was being made only to the 'typical immigrant' in this study.

That needs which were satisfied in Britain but which were not being satisfied in Australia are important to investigate is not disputed. Mention of their potential importance was made on p. 10 of the monograph and several of the findings—e.g. incidence of homesickness—might be interpreted as the result of this type of need being unsatisfied. However, for the group of migrants studied, overall satisfaction with life in Australia appeared to be largely determined by the extent to which the needs that were unsatisfied in Britain were fulfilled in Australia. In considering the

satisfaction scale it is important to realize that it was designed to measure overall satisfaction with life in Australia rather than specific satisfactions and, consequently, was equally applicable to any immigrant whether his reasons for emigrating were based on material or non-material 'value orientations' or whether his present situation and experience in Australia had provided additional satisfaction, or dissatisfaction that had not been anticipated. Only the so-called validation criteria for the satisfaction subscale were selected from some of the known satisfactions sought by the majority of immigrants of this particular socio-economic level (see Richardson, 1959). Dr Emery evidently misunderstood this fact as he complains that the satisfaction scale consists of 'an admixture of conceptually unrelated items that deal with specific grievances'. Of the six items on this scale, four deal directly with overall satisfaction with life in Australia, one with district satisfaction and one with accommodation satisfaction. Only this last could be thought to deal with a specific grievance.

Dr Emery, in a brief reference to the identification subscale, thinks that it is based on a 'complicated and ambiguous operational definition' and that 'practically all the work of the scale could have been done by one item'. Three measures were included with the aim of investigating three different aspects of identification with Australia. The three measures were found to fit into an acceptable cumulative scale and a cutting point was decided when scores on this scale were being examined in relation to the formation of the overall assimilation scale. It is unclear why this standard procedure should be considered either 'complicated' or 'ambiguous' or how a decision on the discriminating power of any item relative to the assimilation scale could have been decided beforehand.

By accommodation is usually meant a process of adjustment in which outward changes in conventional modes of behaviour are adopted with the object of reducing unnecessary conflict with the host culture but with no assumptions as to the occurrence of any inner changes in beliefs, attitudes or values. People more or less rapidly accommodate to new situations but acculturation usually takes longer because of the implied inner changes. However, in discussing the acculturation subscale Dr Emery believes that it is 'conceptually more relevant to the notion of accommodation than to acculturation', despite the fact that scores on the acculturation subscale are significantly higher for those who have been in Australia for six or more years compared with those who have been in Australia for less than five years. Dr Emery is also in error in believing that, 'the so-called Australian Norms Test essentially deals with whether migrants have picked up the local slang and dropped common British beliefs . . .'. The Australian Norms Test is a completely separate measure of acculturation from the Slang Usage Test and is not restricted to the measurement of dropped British beliefs but is equally concerned with the development of new beliefs that are only relevant within the local Australian situation, e.g., 'It is sensible for parents to allow boys to go to school barefoot in the summer months.'

In conclusion a word must be said on the atypicality of the community within which the immigrants and Australians in this study were living at the time of interviewing. Dr Emery believes that 'The reader might too easily forget just how

atypical this community is'. This seems a curious and gratuitous comment when it is realized that in the monograph under review, after describing the community, a specific warning is given that '... it should be borne in mind that the community is in many ways an atypical one. It cannot be sufficiently emphasized that absolute percentages cannot be generalized with any degree of confidence to the situation of Australians and British immigrants in Australia as a whole' (Richardson, 1961, p. 7), and again, on page 9, 'The results of this study are based on a particular sample of British immigrants at a particular time and in a particular place. Generalization from the substantive findings is dangerous. Further studies using the assimilation scale with other samples of immigrants including non-British immigrants are required.'

In this regard it is of interest that the conceptual scheme and the measuring instruments have received further support from a replication study of Dutch immigrants in Newtown (Taft, 1961). However, in a somewhat different study using other types of measures (Taft and Doczy, 1962), it was found that the assimilation of Hungarian intellectuals may follow a different pattern of assimilation. The time has not yet been reached when any single theory can hope to embrace all the facts, and it is for this reason that the insightful observations made by Dr Emery on other factors related to the assimilation process may assist in the ultimate development of a more elaborate conceptual scheme for integrating the widely varying behaviours of different immigrant groups.

REFERENCES

EMERY, F. E. (1962). Critical notice. *Brit. J. soc. clin. Psychol.* 1, 1, 72-75.

RICHARDSON, A. (1959). Some psycho-social aspects of British emigration to Australia. *Brit. J. Sociol.* 10, 4, 327-337.

RICHARDSON, A. (1961). *The Assimilation of British Immigrants in a Western Australian Community*. R.E.M.P. Bulletin 9 (No. 1/2), pp. 75.

TAFT, R. (1961). The assimilation of Dutch male immigrants in a Western Australian community. *Hum. Relat.* 14, 3, 265-281.

TAFT, R. & DOCZY, A. G. (1962). *The Assimilation of Intellectual Refugees in Western Australia: With Special Reference to Hungarians*. R.E.M.P. Bulletin 9 and 10 (No. 1/2), pp. 82.

Manuscript received 15 August 1962

Book Reviews

Creativity and Intelligence: Explorations with Gifted Students. By Jacob W. Getzels and Philip W. Jackson. London and New York. John Wiley & Sons, 1962. Pp. xvii + 293. 49s.

This book is concerned mainly with the differences between the highly 'intelligent' and the highly 'creative' American adolescent. The qualification 'highly' is necessary since the majority of the (450) subjects were intellectually well above average, having a mean I.Q. of 132 and being drawn from a private school of very high academic standing. The authors, both educational psychologists at the University of Chicago, unashamedly use 'intelligent' to signify high I.Q. (on Binet or Henmon-Nelson) and 'creative' to signify high scores on a battery of 'creative tests'. What most of these tests had in common was that the score depended not on a single predetermined correct response, as is most often the case with the common intelligence test, but on the number, novelty, and variety of adaptive responses to a given stimulus task. These tests included, for instance, 'varied associations to stimulus words; the ability to structure quickly an incomplete or distorted perceptual stimulus; the ability to see numerous different problems in a single set of numerical data; remote, or clever, or original responses to complex verbal situations.' Despite the circularity of this definition of 'creativity', and the inevitable subjectivity of some of the scoring, many of the results gained are interesting, fruitful and not wholly predictable.

The 'high creativity group' consisted of (26) boys and girls who came in the top 20 per cent on the summated creativity measures but in the bottom 80 per cent on I.Q. The 'high intelligence group' consisted of (28) boys and girls who came in the top 20 per cent on I.Q. but in the bottom 80 per cent on the creativity measures. Most of the book is concerned with the differences found in the personalities, the home backgrounds and the social and scholastic records of these two groups. The rationale for this procedure is based on the overemphasis, in the authors' view, that has been placed on 'the I.Q. metric' in relation to the gifted child. They point out that the word 'gifted' has, in the literature, become virtually synonymous with 'high I.Q.'; that the traditional intelligence test 'represents a rather narrow band of intellectual tasks, relying chiefly on those requiring in Guilford's terms "convergent" as opposed to "divergent" thinking'; that the correlation between learning and I.Q.—although the latter is 'probably the best single measure we have—rarely accounts for more than one-quarter of the variance in such crucial factors as school achievement and academic performance'; and that 'the I.Q. metric has been peculiarly immune to advances in our understanding of thinking and behaviour'.

The aims of Getzels and Jackson thus seem to be well worthwhile and their results largely justify their methods. They find, for instance, that the 'high creativity group' do as well scholastically as the 'high intelligence group', but that the former tend to be less well liked by teachers and less success-oriented generally than the latter; that the 'high creativity group' are more humorous and 'intellectually playful', more relaxed, less conforming, socially and intellectually, and more aggressive (or more willing to admit to aggression) in their fantasies than are the 'highly intelligent'; that the parents of the 'high creatives' tend to be less vigilant, less 'bookish' (i.e. magazine-taking), less obsessional and critical of their children than the parents of the high I.Q.s.

These findings, based on extensive questionnaires and interviews, accord well with common sense and observation and introspection. The authors then try to fit their results into a 'theoretical and educational context'. In the first of these aims they are less successful. They consider in turn (a) logical, associationist and Gestalt conceptions, (b) psychoanalytic and neo-psychoanalytic conceptions, (c) perceptual and social conceptions. Their treatment of (a) and (c) strike the reviewer as weak, and somewhat twisted, in order to fit into a not quite relevant framework. The section on psychoanalysis, however, is interesting, with its

stress on the role of the preconscious as opposed to the unconscious, in 'creative' and 'non-creative' work. On the 'educational context' aspect, Getzels and Jackson show insight and courage, daring to defend to a primarily American readership, the value of 'periods of frustration'; the encouragement of able students to pursue their own interests 'even if it means working alone'; the choice of an 'unrealistic' career; and even non-objective, long-answer, essay-type tests.

The remaining chapters are rather disappointing. In one, the highly 'moral' is compared with the highly 'adjusted' adolescent. Here, the acceptance of questionnaire data as the criterion for 'morality' is questionable, as is the relevance of this chapter to a book about creativity. The last chapter, 'Clinical Studies', should be one of the most interesting but, surprisingly, lacks vividness and adds little to what the reader has already learned. The Appendix lists and describes the various tests and questionnaires used in the investigation.

Minor criticisms of the book include an inadequate index; an infuriating method of arranging references; and lack of discussion of the children who were both highly 'creative' and highly 'intelligent'. More important are some indications of being wise after the event (judging from some of the excerpts from 'creative' and 'non-creative' subjects' writing) and of occasional unwillingness to go more than half-way (e.g. 'the outcome [of the creatives] was inventive, not bizarre'). In general, however, it is a stimulating book on a worthwhile subject. It should certainly be read by all psychometrists and all anti-psychometrists.

A. W. HEIM

Mental Health in the Metropolis: the Midtown Manhattan Study, Volume 1. By Leo Srole, Thomas S. Langner, Stanley T. Michael, Marvin K. Opler and Thomas A. C. Rennie. McGraw Hill Book Company Inc., New York-Toronto-London, 1962. Pp. 428. 77s.

In 1952 an interdisciplinary team of psychiatrists, psychologists, sociologists and anthropologists, under the leadership of the late Thomas Rennie, began a monumental study to investigate the consequences of socio-cultural conditions on the mental health of a residential area within Manhattan. The results of this vast enterprise will be reported in three volumes; Volume 1 describes the relation between mental health and demographic factors; Volume 2 will be concerned with the relation of family structure and function (the authors call these 'component' variables—why?) and mental health; Volume 3 will present more intensive explorations of Midtown's ethnic families. The present volume, as the first in the series, also provides an exposition of the concepts and ideas underlying the study, and of the methods used in the vast collection of data. Throughout the volume, ideas, methods and findings are confronted with those from other epidemiological studies, so that, it incidentally, presents also a review of the major current approaches to questions of social psychiatry. Incidentally, too, the authors have provided a demographic picture of Manhattan which throws some light on the reasons for the peculiar loneliness of people in that great city.

In their central contribution the authors start with the assumption that mental health and mental illness have diverse roots; their search for pathogenic environmental factors which may intensify, predispose or precipitate mental illness does not lead them to a denial of the importance of organic factors; here they are in line with all good epidemiological research. There is, in the volume, noticeable particularly in the epilogue, some evidence to show that this interdisciplinary team faced intense intellectual struggle between the psychiatrists and the social scientists before they clarified their position on the constitutional and perhaps hereditary component in mental illness. As a result, the volume is exemplary in the caution of interpretation, the consideration of alternative hypotheses, and the general tendency to qualify and question the results as they emerged.

The data were collected in two basic ways: a treatment census was conducted in the area,

including public and private hospitals, clinics and individual practitioners to assess the prevalence of mental and emotional disturbance in the area. But since the number of persons in treatment at any given time is inevitably an underestimate of the actual morbidity in the population, the authors' main concern is with their second source of data, a home interview survey of an area sample consisting of 1600 persons between 20 and 59 years of age. While 13 per cent of the so selected persons refused to co-operate, some evidence is presented to indicate that as far as the independent demographic variables are concerned the actual sample is unbiased. The interviews, which lasted on the average 2 hours, covered a wide range of topics: physical health, psychosomatic conditions, childhood experiences, interpersonal relations, fears, conflicts and psychological complaints, family composition (past and present), mobility, religion and religious practices, ethnic background, socio-economic status, etc.

From this information, the dependent variable, a psychiatric rating scale of degree of symptom formation, was obtained. Rating in six categories ranging from 'well' to 'incapacitated' was performed from the questionnaires by psychiatrists who rightly refrained from specific diagnoses but were concerned only with the severity of symptoms. Anyone interested in methods of rating and the problem of exclusion of bias will profit much from reading the account of procedures in this respect.

By this standard with its acknowledged and inevitable limitations 23.4 per cent of the Midtown population was found to be psychologically impaired (a combination of the lowest three scale points: marked symptom formation, severe symptom formation and incapacitated); only 1 in 20 of these impaired persons was actually receiving treatment. The proportion of well persons was smaller among those above 30 years of age than among the younger adults, lowest among single men and the divorced of both sexes. Impairment is more frequent among those whose parents' socio-economic status was low; this relationship is even stronger when a person's own socio-economic status is considered. There is a particularly sophisticated analysis of the mental health of immigrants, taking the socio-economic background in the home country into consideration and producing some evidence to the effect that impairment proportions are highest among recent immigrants, lower in subsequent American born generations; and an interesting analysis of religious affiliation and mental health.

These few examples of research results should indicate the scope and value of the study: it is a painstaking effort in an immensely complex field to weld social science methods and psychiatric concepts into a tool which may one day help to unravel the etiology of mental illness. It is a book for the expert rather than the layman, who will welcome it even though it does not contain any startling new insights; for the experts know that with regard to mental health and mental illness solid scientific craftsmanship, as demonstrated in this book, is scarcer and more urgently needed than apparently brilliant but unfounded generalizations.

MARIE JAHODA

The Meaning and Measurement of Neuroticism and Anxiety. By R. B. Cattell and I. H. Scheier. The Ronald Press Co., 1961. Pp. 468. 92s. 6d.

We are told that this book is directed toward two kinds of people, the clinician and the researcher. The authors say that their work will assist the former group by increasing the reliability of diagnosis and treatment by using and understanding precise scientific instruments to measure levels of anxiety and severity of neurosis, while those interested in research will be given a 'firmer basis of empirical findings about personality structure'.

In their preface to this book the authors also hold out the hope to the reader that he will find in it '... conceptually clear and precise instruments of measurement for use in research and routine practice'. However, we are warned that the proper import of the techniques and procedures referred to can only be grasped by the reader who is prepared to accept radically new ideas.

In the reviewer's opinion the book falls far short of the achievements confidently referred to in the preface. In fact the 'new ideas' we are asked to grasp turn out to be time-worn, and the new techniques and procedures amount to nothing more than the multivariate analysis of old and new data gathered by Cattell and other workers—the principle of selection employed in the case of the latter being undisclosed.

As for the 'precise scientific instruments', these are, it seems, questionnaires and objective tests apparently characterized by poor reliability and doubtful validity, and of virtually unknown clinical value.

One is also left with the impression that an adequate presentation, or even in some cases the bare mention of a particular piece of information would have greatly contributed to the meaningfulness of this book. Sometimes the compression of material has resulted in near incomprehensibility, a feature which is not lessened by Cattell's predilection for neologisms in labelling his factors.

However, the evaluation of this work really hinges upon two main points; the procedures used and other technical matters related to the statistical analysis, and the contribution made by this book to clinical understanding and know-how.

So far as the first of these points is concerned Cattell is open to criticism for precisely those reasons given in papers by Sir Godfrey Thomson and Dr Charlotte Banks. Indeed, the discussion of Cattell's earlier contributions by these two authors seems to be as apt today as they were fifteen years ago, and suggest that his work has a less secure statistical foundation than one is asked to believe.

It is, however, the second main point, concerning the clinical value of the book, which most readers will be anxious to appreciate. Certainly Cattell will carry all clinicians with him in stating that there is a '... need for more effective treatment and more adequate theories, a genuine dissatisfaction and search for new possibilities ...', but it is unlikely that many readers will endorse the solution offered. For example, it seems reasonable to ask questions concerning the way in which an individual's factor profile is obtained. Basically this depends upon the satisfactory nature of the 16 PF questionnaire, although the authors refer to the Objective-Analytic battery of tests which is apparently being revised to improve its validity.

In general the position respecting personality questionnaires of a conventional character seems not to have changed much in recent years. Kornhauser (1945) sounded the opinions of psychologists on questionnaires and discovered, not surprisingly, that 85 per cent of the sample considered such procedures doubtfully satisfactory or worse. Cattell's data does little or nothing by way of converting this large majority. The reliability coefficients are not at all comparable with those obtained from cognitive tests, and indeed in the 1957 IPAT manual no information is given concerning test-retest reliabilities because certain traits 'fluctuated in level a good deal'. This may be the case but such a conclusion must necessarily limit the clinical usefulness of the knowledge derived from the test.

Further difficulties might be envisaged by the more cautious clinician in adopting the procedure of correlating an individual's profile with the 'standard profile' obtained by various diagnostic groupings. In effect this would mean offering the statement that patient x is obsessive-compulsive to the extent of r_{0-78} etc. This technique could not only be faulted on the grounds that the unreliability is being multiplied, but that the end result is tied up with what are commonly regarded as unsatisfactory diagnostic categories and that the profile itself has no obvious implications for treatment and disposal.

Perhaps just as important an issue is the degree of differentiation between groups on the various factors. Here the criticisms of the interested reader of this book might range from doubting the logic of some of the arguments advanced to astonishment at the conclusions drawn from the data provided.

But in justice to Cattell and Scheier it must be said that a note of caution enters the discussion from time to time, indeed, in treating of diagnosis and therapy it is admitted that half of what is being said is only a statement of things to come. It must also be said that bringing together the mass of data presented represents an enormous task, and clearly a great deal of work has gone into the preparation of the book. However, clinicians and researchers alike are likely to be somewhat sceptical about the procedures employed, the conclusions drawn, and the rather extravagant claims made for the contents of the book.

It is probable that the volume could be re-written and carefully edited to eliminate obscurities and present a more balanced evaluation of the contribution made by this material; it is more difficult to see how the techniques and procedures used in obtaining the material could be improved by such a revision.

H. R. BEECH

REFERENCE

KORNAUSER, A. (1945). Replies of psychologists to a short questionnaire on mental test development, personality inventories, and the Rorschach Test. *Educ. Psychol. Measmt.* 5, 3-15.

A Textbook of Psychosexual Disorders. Clifford Allen. London: Oxford University Press, 1962. Pp. 408. 42s.

'In this book an attempt has been made to examine sexual abnormality in the same way as any other illness.' To this end the author outlines instinctual development in animals and man, followed by brief mention of genetic and endocrine factors and a much more extensive outline of psychoanalytic theory. The sexual disorders, classified by a combination of their mode of expression and sexual object, are then considered seriatim, with final chapters on prevention, treatment and prognosis. His approach is eclectic but predominantly psychoanalytical. His message is that 'psychosexual disorders are as capable of treatment as any other neuroses'. Of 17 illustrative cases stated to have been treated, 15 were successful, one partially so, and one unsuccessful.

There is much to be learnt from the book, not least the need for a compassionate and optimistic approach. Nevertheless it is disappointing in some respects. Terminologically, the paraphilias, psychosexual disorders, perversions and sexual neuroses are used sometimes particularly and sometimes interchangeably; sadism, psychopathy and schizophrenia are ascribed respectively to agenesis, dysgenesis and over-severity, of the superego. His conception of 'obsessional compulsion' (p. 108) would include all the appetites, normal or otherwise. We must take or leave his statement that the homosexuals love object is invariably the mother in disguise; this may be true of everybody, or is this irrelevant? 'Results are infinitely better when the treatment is given privately than when given at clinics'; but what of the selection of patients? Many key references are unmentioned: Klüver and Bucy, and Van Reeth on hypersexuality; Lucianowicz on transvestism; Bremer on castration; Glover on prostitution, Gibbens on their clients; Masserman on masochism; Epstein on fetishism. The section on medico-legal aspects is out of date: no mention of the Mental Health Act, and repealed acts still included; there have been two editions of Kenny since the one he lists.

Aversion therapists may be interested to know that Basil James' double approach, of conditioning followed by encouragement, was used by Srnec and Freud in 1953.

P. D. SCOTT

Ermüdung, Beanspruchung und Leistung. Dr. Manfred Haider. Vienna: Franz Deuticke, 1962. Pp. 168. Öst. S. 158.

This is a short, well written and comprehensive account of experimental work in the ill-defined and very controversial field of fatigue. The writer starts off with a review of experiments done in industry, goes on to a consideration of 'tests of fatigue', considers physiological factors, as well as monotony and satiation, and then continues with more theoretical concepts such as psychological criteria of fatigue, clinical work and a consideration of the reticular activating system. There is nothing startlingly new in this volume, but it is extremely comprehensive, considering its shortness, and unlike so many continental writers

the author is very well acquainted with the relevant research literature in English. (This is perhaps less surprising when it is realized that he has actually been engaged on work along these lines in this country.) As an introductory text this one would take a lot of beating, and it can be wholeheartedly recommended to psychologists interested in this field. Even those who feel that they know the literature well will be surprised here and now to see references to early German work, particularly of the Kraepelin school, which antedates subsequent better known American work. Thus, for instance, Bills on blocking was anticipated by Voss who experimented along similar lines before the turn of the century.

One very important omission considerably reduces the value of the book. Nowhere does the author deal with the Hullian concept of inhibition, which is perhaps the best known psychological theory linking up the many experiments he quotes; indeed he does not even cite work on performance decrement carried out by psychologists working within this general framework. This omission of a large and obviously relevant area can hardly be accidental, yet the reviewer is at a loss to explain its intentional omission. All that can be done is to note it here and hope that in future editions a little more will be said about the omitted portions, both on the theoretical and on the empirical side.

H. J. EYSENCK

Society: Problems and Methods of Study. A. T. Welford, Michael Argyle, D. V. Glass & J. N. Morris (eds.), London: Routledge & Kegan Paul Ltd. Pp. 586. 50s.

Rather more than twenty years ago a symposium entitled *The Study of Society* enjoyed a considerable and well merited success. Its object was to 'state clearly what psychology, anthropology and sociology have done and can do . . . to aid in the solution of pressing social problems'. The book under review seems to be an attempt to bring the earlier volume up to date although, naturally enough, changed circumstances and new editors have produced some interesting differences. The main emphasis now is on sociology and social psychology and concern with social problems is somewhat oblique and generally less in evidence. Anthropology, which figured so prominently in the earlier volume, now disappears entirely and apart from interesting chapters by Asa Briggs and Bryan R. Wilson on, respectively, sociology and history, and the analytical study of social institutions, there is nothing on those historical and philosophical aspects of sociology which are prominent both in the Hobhouse-Ginsberg and in the general European tradition. The temper of the book is, in fact, largely that of scientific sociology and psychology with the emphasis placed upon empiricism and quantification. Of course no symposium can cover everything and the book is undoubtedly better for being focused in this way. But even within these limits it is surprising to find nothing on social perception, race relations, community studies and the sociology and psychology of religion while space is found for a chapter on personnel selection and for two chapters on accidents.

A book with thirty-one chapters and rather more contributors is bound to be rather uneven and 'Society' is no exception, although the general standard is very high indeed and each chapter contains at least something of value. O. R. McGregor and Griselda Rowntree contribute a chapter on the family which provides a useful summary of some of the relevant demographic data and which is also well worth reading for its astringent wit and literary style. Those critics who complain of the inelegance of much sociological writing might well be referred to this chapter as evidence that competence in the subject and an inability to write are not necessarily related. A. T. Welford contributes a chapter on experimental psychology in which he shows the relevance of this subject for an understanding of social behaviour and he finishes with an eminently sensible plea for more co-operation between the experimental psychologist and the sociologist. Jean Floud writes on educational sociology, P. E. Vernon on the measurement of abilities, attitudes and personality, Michael Argyle on small group research while B. Benjamin contributes a useful chapter on the British population census. There are, in addition, many other excellent things too numerous to mention in a short review.

It is safe to predict that 'Society' will be as successful as its predecessor and the editors are to be congratulated on corralling an excellent team of contributors who between them have produced a volume which, within its limits, will be of value both to students and to the lay public.

PETER COLLISON

Work and Leisure. By Nels Anderson. London: Routledge, 1961. 28s.

At first sight, nothing could seem simpler than the distinction between work and leisure. Work is paid activity, in which the worker is under a definite social obligation. Leisure, in contrast, is spare time—unpaid and free from obligations. Seen from this point of view, work and leisure are clearly distinguishable ways of using time. Together, they seem to cover the whole of the waking day.

On reflection, the distinction raises difficulties. Do work and leisure necessarily exclude one another as ways of using time? What of Mark Abrams' claim that many people, and particularly women who are hungry for day-time companionship, regard work as a leisure-time activity? Even if we agree to regard the two activities as separate, are we right in linking obligation so closely with payment? Maybe for most of us the sinecure is dead, and our paid work involves us in obligations. But are obligations always paid for? What of the thousands of unpaid voluntary workers who form a vital part of our social services?

Professor Anderson, who is Director of Research of the Unesco Institute for Social Research, has taken up these issues, and very many others, in a brisk, comprehensive and reasonably brief study of the links between work and leisure. His thesis, first developed in his earlier book *The Urban Community*, is that the mass leisure so characteristic of modern industrial society is a by-product of the single-minded concentration on efficiency in work. Leisure, then, has been earned by high productivity, with its associated specialization, mobility and urbanism. Part of the cost of this widely distributed 'spare time' is a sharp contrast between the different parts of one's life, occasioned by the very specific demands of work, which are not necessarily related to personal needs or interests. A great deal of our social life is sympathetically viewed as being a series of attempts, many of them quite successful, to achieve a 'rationale', or sense of personal integration. Work, for example, becomes a source of relatively mild interest which is limited to the work situation itself. Stronger interests are found in leisure, of which there are two kinds, 'leisure as play' and 'non-work obligations'. The first kind of leisure is concerned with individual and collective games and pastimes. Professor Anderson has little time for 'viewing with alarm' and believes that in the long run the amount of participation will increase (as it is doing), and that the high level of some professional sports and cultural activities amply justifies large numbers of spectators. The second kind of leisure is of very great importance in our complex and highly differentiated society, since many important social functions in the family, neighbourhood and community are associated with these unpaid obligations, such as looking after the family and home, assisting neighbours, doing voluntary social work, taking part in political and religious activities, and civic obligations of every kind. For many people, these obligations, often chosen with greater freedom than one's work and often more easily changed, displace 'leisure as play' almost entirely, and without loss of satisfaction. There is clearly a great deal of room for study of these two kinds of leisure, and Professor Anderson tries to show us some of the connections with Western ideologies of work and play.

Three of the book's ten chapters are given to a study of work and leisure in three main stages of the 'life-cycle'—youth, maturity and old age. A sustained attempt is made to bring in relevant empirical studies but the treatment is understandably uneven. As the author reminds us, the study of work and leisure is still very undeveloped. Some of the unevenness, however, is his own. His characteristic blend of comprehensiveness, brisk generalization and optimism often leads to a lack of rigour in handling his extremely varied materials.

One more carping comment and I can close on a more positive note. In a book which is clearly meant for students, it is important to provide the clearest and simplest ways of following up references. *Work and Leisure* is lacking in two respects: it has no bibliography, although there is a source reference on almost every page, and there is hardly any mention

of the inclusion of references in relevant books of readings. For example, no less than seventeen of the references made to original sources (sometimes in out of the way journals and books) are to be found reprinted in Larrabee and Meyersohn's *Mass Leisure*. Obviously there was no obligation on the author's part to mention this, but diligent readers would have gained more leisure had he done so.

This said, I must add my view that *Work and Leisure* is the obvious introduction to a topic of great and growing importance.

J. F. MORRIS

Mirror of Minds. By Geoffrey Bullough. University of London: The Athlone Press. Pp. vi + 271. 35s.

Relating literature to 'the History of Ideas' is far from a new gambit, but it still has much to offer, and Professor Bullough investigates a particularly fruitful field. The creative writer and the psychologist have always been close—inevitably, since both are concerned ultimately with the character of Man. *Mirror of Minds* is concerned with the influence (not always consciously realized) of psychological theory on certain writers; another book might well be written on the debt of psychological theorists to literature—Freud, after all, credited the discovery of the unconscious mind to the poets. Professor Bullough has had to limit himself drastically in his Alexander lectures. His attention is almost exclusively confined to poetry; the drama is not mentioned, apart from some pages on the Morality Play and on Shakespeare, and the novel is only noticed from time to time, and then briefly. One has no right to complain, and Professor Bullough might well say that he has covered a great deal of ground as it is. The subject is a vast one, and while it is comparatively easy to give a coherent account of Medieval and Renaissance psychological theory and illustrate it from literature, it is much harder to cope with the enlightenment and almost impossible to offer an intelligible synoptic account of the interaction of Psychology and Art in the last hundred years. *Mirror for Minds* is full of information and only sometimes does it seem to lose its way, or degenerate into a catalogue. The limitation to poetry, however, does produce a sense of unbalance, and it is unsatisfying, for example, to find reference to Joyce limited to one sentence. At the same time Professor Bullough occasionally seems to get outside the proper limits of his subject: is impressionism, which he treats, really psychological theory? Perhaps, stimulated by these lectures, one is really asking Professor Bullough to write another book. One howler must be noticed: on p. 72 Melancholy is described as 'the hot dry humour', a definition which would much have surprised Robert Burton, to whom it appears to be credited.

J. B. BAMBOROUGH

The Control of the Mind. A Symposium edited by Seymour M. Farber and Roger H. L. Wilson. New York: McGraw-Hill. Pp. xvi + 340. Paper back 23s. Cloth-bound 50s 6d.

In January 1961 a somewhat heterogeneous collection of experts met at the University of California San Francisco Medical Center to discuss 'Man and Civilization: Control of the Mind'. This book is a record of the proceedings. The broad title accounts for the variety of topics actually discussed, though the majority of the papers were concerned with the narrower issue: the control of the mind. There is also an indication that the opportunity was taken to bridge the gulf between the students of the arts and those dedicated to the pursuit of science. Hence the presence of Father D'Arcy, Arthur Koestler, and the historian Professor H. Stuart Hughes.

So far as the main topic is concerned, many of the contributions are extremely informative, particularly those dealing with the biochemical properties of the brain and the effect of drugs on the nervous system. This part, which comes first, was conducted by the physiologists and the pharmacologists. What is known of the biochemical processes in the central

nervous system is described and we learn, for example, that protein synthesis 'represents today the most widely accepted chemical process by which information can be stored by the brain', and that 'if the genetic information is stored in chemical structures or physical arrangements of chemical molecules, then chemical or physical forces can change these patterns'.

Granted that chemical processes play a large part in the working of the brain, it is reasonable to expect that their activities can be influenced by drugs. The pharmacologists, however, took the line that they can only affect behaviour temporarily, that they can repair damage but cannot improve normal performance, and that owing to their limited action and also to idiosyncratic differences in response, they could not be used effectively in any large-scale attempt to control the minds of the public at large.

Later on the influence of the press and other forms of communication were discussed. Professor Porter, Professor of Journalism at Iowa, expressed the view that the influence of mass-media is very much exaggerated. This position was agreed to by Professor Lasswell. He, however, was concerned with a broader issue. In his opinion mass-media tend to reinforce parochialism—national differences and the like. This he deplores, but he suggested that there are two universal activities of the mind—scientific modes of thought and the appreciation of art, and that in so far as these become universalized by widespread communication parochialism will be undermined.

So far as control of the mind is concerned, the only effective methods are the negative one of withholding news or misrepresenting it without the means of correction being available, and the psychological technique of alternating fear and promise of reward, which is alleged to be used on individuals by the Chinese. Even in this latter case Father D'Arcy appears to hold that certain regions of the person are unassailable.

So much for the contributions which deal directly with the main theme. In addition Professor Hebb contributed a paper on the need for constant stimulation if the mind is to work properly, and the concomitant need for social order if the responses are not to get out of hand. Aldous Huxley talked informatively about the possibilities of improving our mental powers and enjoyment, quoting, as one would expect, from little known and unorthodox literature. Professor A. C. Mace considered the problem of motivation in affluent society, when we are no longer driven by biological needs. In his paper he criticized those views which imply that living organisms only 'live' between periods of biological satiation; if this were the case the domestic cat and dog, who live in affluence, would spend the greater part of their lives curled up asleep. One of the main non-biological motives in man, as he sees it, is the enjoyment of work well done. His view of the future is engaging: 'The skilled workers will be perfectionists in their work; the displaced unskilled workers, on the other hand, will become perfectionists in their innocent recreational pursuits, breeding better and better pigeons or greyhounds, cultivating perfect auriculas, becoming better ballroom dancers, or pursuing in a perfectionistic way biologically useless goals, but enjoying in these pursuits the good life'.

Arthur Koestler's paper stands apart from the rest of the contributions, but is none the less worth reading for that. He was concerned with depicting the creative processes in humour, in art, and in science as proceeding from a clash between different and opposed frames of reference: the self-assertive and the self-transcending sides of our natures, which may, he suggests, be related to the contrary functions in the autonomic nervous system.

It cannot be said that the papers, taken together, form a coherent scheme; that would be more than one could expect from such an array of contributors. It is, however, remarkable to have a collection of papers, each of which can be read with pleasure and profit.

W. J. H. SPROTT

Psychological Statistics. By Quinn McNemar. New York: John Wiley and Sons.
3rd Edition, 1962. Pp. vi + 451. 59s.

In a recent review of books recommended by departments of psychology in American universities the esteem in which this book by McNemar is held became clear. It was first published in 1947 and, as is well known, it provides an introductory course in statistics for

students in psychology and education. In this, the third edition, the author has made a number of extensions in an attempt to keep the book up-to-date, the principal of which is a chapter on trend analysis (not very helpful), and the discussion of such topics as *distribution curves, probability and hypothesis testing*, and questions of *inference* is also supplemented.

As McNemar's book is already so well known it would be superfluous to give a detailed list of contents: it contains a concise and clear account in non-mathematical language of most of the elementary statistical techniques in common use. What seems more important in reviewing the book is to look at this new edition in the light of recent developments in statistical theory and practice and see if it can still be relied upon to give a non-specialist student of statistics an enlightened picture of present-day procedures.

In the development of statistical theory there have been three distinct outbursts of activity, one by Karl Pearson, another by R. A. Fisher, and a third by A. Wald and present-day subjective probability theorists. Pearson's work was concerned primarily with descriptive statistics and theoretical distributions: Fisher's main contribution was perhaps in laying the foundations of a scientific theory of experimentation; while Wald, and those influenced by him, are concerned with the basic foundations of statistical theory and the processes whereby we reach decisions. If we omit further discussion of this latter group, who are still battling at the frontiers of knowledge, we are left with the contributions of Pearson and Fisher and the many disciples who develop and are still developing their ideas.

In keeping with Pearsonian ideas McNemar's book is very suitable for the non-specialist. It gives an adequate account of descriptive statistics and elementary distribution theory. Simple tests of significance of the large-sample type for comparing means and proportions also are clearly described. Small sample methods are also introduced and, as might be expected, the book has an excellent account of correlation theory and of methods of comparing, averaging, and testing the significance of correlation coefficients. Similarly it deals admirably with straightforward questions of regression, and with problems of association and 'goodness of fit'. But all this is reminiscent of the days when people first collected their data and afterwards approached a statistician about how to describe and summarize them. Indeed, when one is dealing with large samples which are fairly representative of the populations they are taken to represent, then these routine methods are often all that is needed. But psychological research today is more and more concerned with deliberate experimentation, that is with small efficient experiments carefully planned to test specific hypotheses and theories, rather than with big general investigations of a purely descriptive type. Here the competent worker will need to make himself familiar with the methods of experimental design initiated by Fisher. It is with regard to this aspect of statistics that the present book is outdated, as compared with a book such as *Methods of Statistical Analysis*, by C. H. Goulden. True, there are carefully written chapters on analysis of variance and covariance, which explain how a composite estimate of variance can be partitioned into components due to different aspects of the investigation. Variance ratio tests are described and the assumptions which underly them are clearly enumerated, yet essential information on how investigations can be planned and on the most efficient ways of collecting the data in the first instance are missing. It is not sufficient to retort that this is not a book on the design of experiments but one on statistical techniques, for the two aspects of experimental procedure can no longer be separated. McNemar's book is likely to continue to be useful as a reference book on techniques but it could not be relied upon to convey to our students the true spirit of present-day statistical procedure.

A. E. MAXWELL

A Therapeutic Experiment on Phobic and Affective Symptoms in an Individual Psychiatric Patient

By M. B. SHAPIRO, I. M. MARKS AND B. FOX
Institute of Psychiatry, Maudsley Hospital

An experiment was carried out on the symptoms of a 38-year-old female patient with phobias and depressive symptoms of long standing. Two forms of treatment, one 'rational-training' and the other 'non-directive', were applied to the patient. Eight sessions of each type of treatment were completed over a period of four weeks, in balanced order, with measures of both immediate and delayed effects of each type of treatment session upon the symptoms. The main findings were: (i) each form of treatment produced beneficial immediate effects; (ii) the delayed effect of the 'rational-training' sessions appeared to be that of relapse, and that of the 'non-directive' to be variable though with an average effect of improvement; (iii) previous findings of the relatively independent fluctuations of phobic and depressive symptoms were confirmed.

INTRODUCTION

It has been argued elsewhere (Shapiro, 1961a, 1962) that there is a need in clinical psychology for a new type of research which (i) is consistently and systematically based upon individual patients, (ii) is directed at those psychological dysfunctions which are primarily responsible for the patient's presence in hospital, and (iii) consists of the systematic experimental manipulation of psychological variables which have therapeutic relevance.

We know of only one example (Shapiro and Ravenette, 1959) of such an experiment. The subject, in that experiment, had paranoid symptoms. The independent variable consisted of sixteen discussions, of which eight exposed the palpable falseness of the paranoid beliefs, and eight were concerned with coming to terms with guilt feelings. The dependent variables, the false beliefs, were measured with a questionnaire based on the patient's own statements. The two kinds of discussions were held in balanced order, over a period of four weeks, at the rate of four discussions per week. The questionnaire was applied sixteen times, eight times before interview and eight times after, in a balanced order, and evenly distributed among the two treatments. The results suggested that, while the interviews appeared to have no immediate effects, psychologically meaningful changes took place over the weekends during which no interviews were held. The subject's paranoid beliefs became more intensely held after the discussion of the guilt feelings, and less intensely held after the discussion of the paranoid beliefs. The present paper reports the results of an analogous experiment.

THE SUBJECT AND HER PREVIOUS TREATMENT

The subject was a 38-year-old housewife initially referred for incapacitating phobias of eight years' duration, with an estimated I.Q. of 95 on the Mill Hill Vocabulary Test and 121 on Raven's Matrices. Her father, a maintenance engineer, had died thirteen years previously. He had been a beloved 'Little God' who favoured the

patient. The mother, a submissive woman by whom the patient felt rejected, now lived alone. *S* was the fourth child in a family of six girls. When she was 6, a 15-month-old sister died, and when she was 22 her eldest sister died of eclampsia.

She had been a 'tomboy', with male playmates, and at puberty accepted the female role reluctantly. She had worked as a waitress, shop assistant, and in the WRAF until her first pregnancy. She married, at 23, a man five years her senior. She had played with him when aged 8 and 'gone steady' with him since her 18th year. He was the illegitimate son of a doting mother who initially disapproved of the marriage. Sexual activity, after initial difficulties, was satisfactory. The husband constantly sought demonstrations of his wife's affection which she frequently withheld, with resulting quarrels. They had no friends. Their one daughter, aged 14, was more attached to the father. The husband was at sea between the patient's 27th and 29th years, and he had been a prison warder for the last seven years.

At 30, *S* had a cervical polyp removed and at 31 had a miscarriage at two months. Two days later she became tense and anxious when alone, especially outside her home. These symptoms intensified until she became completely restricted to the home after a move. About this time she ceased quarrelling with her husband and new symptoms developed. (See Table 1.)

S was admitted to hospital in October five years later and given psychotherapy for two hours per week by a then inexperienced registrar, I.M.M. From the beginning *S* showed marked fear of her aggressive feelings towards her husband. These were eventually directed at the therapist, in an increasingly permissive atmosphere, and the long history of marital discord began to emerge. Her anxiety appeared to increase when she did not express felt annoyance, and to decrease when she expressed her anger. Two months later I.M.M. pressed her to take walks alone for gradually increasing distances. She laboured hard to overcome great initial anxiety at each new walk. This diminished with practice and, as the patient recognized, lessened if the walk began immediately after an interview, or if she were overtly angry—'I became so angry I forgot my fears'. By February, in the following year, she had lost most of her fears and was going home alone. While she was looking for employment she developed profuse vaginal bleeding. The phobias returned immediately. She had an operation for a cervical polyp in early March. The relapse then was complete, and she became again unable openly to express hostility. I.M.M. suggested that the psychologists might try to treat the phobias with so-called behavioural methods (to be described later) while he continued with a form of non-directive psychotherapy. The possibility of a therapeutic experiment then emerged. The patient had been placed, one week before the experiment began, on a course of Imipramine, at 150 mg. per day, for the duration of the experiment, and after.

PLAN OF EXPERIMENT

It was possible to apply the two kinds of treatment, to be referred to henceforward as 'rational training' or 'T', and 'non-directive' or 'N', systematically for four weeks. There were sixteen sessions of treatment at the rate of four sessions per

week, each treatment lasting for about one hour. Each kind of treatment was given in blocks of two. Thus, the first two sessions were taken up by 'N' treatment, the second two sessions by 'T' treatment and so on. The two kinds of treatment were applied in accordance with the following design: NN TT TT NN NN TT TT NN. The treatments within each block were given on two successive days, except for the second and third psychiatrist's treatments in the first half, for reasons beyond our control. This does not appear to affect the main trend of the results.

The tests (which will be described below) were carried out: (i) immediately after the second of each pair of treatments, i.e. after the even-numbered treatment occasions; and (ii) from one to three days after the second treatment in each block, i.e. immediately before the odd-numbered treatment occasions. There were sixteen measures in all (Nos. 2 to 17 in the Graph and Table 1). For most comparisons, the average rank order of the occasions of testing after 'non-directive' (N) treatment and after 'rational-training' (T) treatment, are the same. (See Tables 1 and 5.) Details of the design of the experiment are given in Table 1.

Table 1. *Plan of measurements*

		First half		Second half	
		Mid-week	Weekend	Mid-week	Weekend
'Non-Directive'	Immediate	2	8	10	16
	Delayed	3	9	11	17
'Rational-Training'	Immediate	6	4	14	12
	Delayed	7	5	15	13

Each number refers to the rank order of the measurement concerned. Measure No. 1 was taken before the experiment began and was excluded from the statistical analysis.

Example: The immediate effects of the 'N' treatments in the first half were measured on occasions 2 and 8, and of the 'T' treatment in the first half, 6 and 4. The average rank order for each treatment is the same: $10/2 = 5.0$.

M.B.S. initiated the 'T' treatment, with T. Mordal, a psychologist, and Miss A. M. Whitney, an occupational therapist; M.B.S. did not participate in most of the training aspect of the treatment, i.e. sessions 11 to 14 inclusive.

MEASUREMENT

Measurements were made with the Personal Questionnaire (Shapiro 1961a, b) which had been administered weekly for clinical purposes ever since I.M.M. began his first treatments. Measure No. 1 in the Graph was the last of these pre-experimental measures. I.M.M. administered the test on all occasions before and after 'non-directive' treatment and Mr Mordal on all occasions before and after 'rational-training' treatment and on the seventeenth occasion.

The Personal Questionnaire is constructed from statements which the subject herself makes in a controlled interview. These statements are then discussed with the registrar and consultant in charge of the case, and a final agreed list, based on a further interview of the patient, is drawn up. Such statements provided the essential content of all the twenty-three statements in Table 2. These statements are formulated to have implications of very great unpleasantness for the patient. Two further statements are constructed for each of the first

statements, thus forming a triad of statements for each symptom. The second member of each triad has moderately or slightly *unpleasant* implications for the *S*, and the third member has moderately or slightly *pleasant* implications. A special procedure, described in the Abbreviated Manual (Shapiro, 1961b), ensures that the patient herself judges the unpleasantness of each of the statements.

Table 2. *Statements of symptoms*

Triad No.	
1.	It is very difficult for me to think of what I want when I go shopping.
2.	When I'm in a shop I often feel I want to run out.
3.	I find it very difficult to find the courage to start my daily walks.
4.	I feel I cannot go into the garden by myself.
5.	I don't like meeting people because I'm lost for words.
6.	I have a fear I may fall down.
7.	I am worried constantly about what effect my state will have on my daughter.
8.	I feel inadequate as a mother.
9.	I feel like a child.
10.	I certainly feel my eyes will give out.
11.	I am afraid I may forget who I am.
12.	My legs feel like jelly.
13.	I get a funny feeling so that I can't think of what I'm doing.
14.	I feel no use at all.
15.	My worries are part of me so that I can't get rid of them.
16.	I feel conscious of my heart's action.
17.	I certainly have a funny feeling in my head and want to poke my fingers into my ears to remove it.
18.	I certainly feel my tummy turning over.
19.	Everything around me seems real and I'm not.
20.	I feel guilty about everything.
21.	I have lost interest in everything.
22.	I feel exhausted.
23.	I feel I've been ill for so long I haven't got a future.

The test takes the form of a set of paired comparisons for each symptom, there being three pairs of statements for each, and each pair being typed on a separate card. There were thus sixty-nine cards for all the twenty-three symptoms included in *S*'s questionnaire. The rationale for scoring is described elsewhere (Shapiro, 1961a, b). Suffice it to say here that there were four main scores possible: 2, 5, 7 and 10, with 2 representing recovery and 10 illness. A test of internal consistency and a method of scoring inconsistent patterns are also provided (Shapiro, 1961b; Phillips, 1961).

The data were acceptable for analysis, as our subject only produced three inconsistent response triads out of a possible 368. These three triads were each arbitrarily assigned a mid-score of 6.

THE TREATMENTS

'Rational-Training' treatment

This was derived, in the main, from Wolpe (1958) and Jones (1956). The treatment went through three phases. The first phase (Session 3 in Graph) consisted of an enquiry into the nature and organization of *S*'s phobias. In the second phase (Sessions 4, 5 and half of 6), a conditioned reflex type of explanation of the phobias, and of the proposed treatment, was expounded to *S* until she was able to state the principles, in spite of an initial diffidence about holding forth in front of authority. She was gently persuaded to inform the therapist whenever she experienced even minimal anxiety. Her feelings of guilt about being ill made persuasion difficult, and the lack of rational and moral justification for such feelings had to be discussed. In addition, arrangements were made to minimize her anxiety outside the treatment situation, e.g. meals were taken in the ward instead of the canteen, which evoked strong anxiety. *S* had to overcome feelings of guilt before accepting these changes.

The third and last phase (half of Session 6 and Sessions 11, 12, 13 and 14 in the Graph and Table 3) consisted of taking *S* for walks when experiencing little or no anxiety, which was inhibited, presumably, by the presence of the therapists. *S* walked for increasing distances, with decreasing contact with the therapists for any particular distance. Walks completed without anxiety were practised. *S* was, at first, almost completely incapacitated. By the last session (Session 14 in the Graph and Table 3) she was able to leave her ward, which was on the second floor, go down the stairs with the therapists out of sight, and walk around the garden (about 150 yards). After two attempts to approach a busy main road with the therapists, and two more circuits of the garden with the therapists in sight, she walked up the stairs back to the ward, with the therapists out of sight. Conversation was always confined to neutral topics.

'Non-Directive' treatment

In marked contrast to the immediately preceding period *S* now talked freely, the themes being aggressive and sexual; the latter being very prominent for the first time. Examples of aggressive feelings were: (i) saying to the therapist, 'At the moment I could really do you a mischief and injury'; (ii) childhood memories of a slaughter house in which pigs were shot in the head, and then had their throats slit, with blood rushing out; and (iii) dreams of her husband being shot, and of herself taking the bullet out of his right side.

Examples of sexual content were: (i) after writing I.M.M. a long note about wanting to discuss sex, spent 2½ hours in frank discussion with husband about sex on 12th June; and (ii) recalling memories such as that of her daughter, when young, saying to the patient's sister 'My daddy calls my mummy "Darling", and then puts his hand up her leg'.

METHOD OF ANALYSING THE DATA

We carried out a factorial analysis of variance (Maxwell, 1958) of the measurements made on occasions 2 to 17 inclusive. We were interested in five kinds of differences and their interactions. They were as follows:

- (i) Kinds of treatment (to be indicated henceforward by the letter 'K') which is of central importance. The actual occasions of measurement involved in each of the two kinds of treatment are given in Table 1 and the Graph. They were, as noted above, balanced for order.
- (ii) The immediate and delayed effects of treatment ('L') (see Table 1 and Graph). The average rank order is 9.0 for the former and 10.0 for the latter.
- (iii) The first and the second halves of the experiment ('H') (Table 1 and Graph): The average rank orders being 5.5 and 13.5 respectively.
- (iv) Weekend and mid-week measures ('W' Table 1 and Graph): The former were made immediately before or immediately after the weekend and the latter were made mid-week because of indications that *S*'s clinical state varied with part of week. The average rank order of occasions of testing was 8.5 for weekend measures, and 10.5 for midweek measures.
- (v) Differences between the 23 triads of the questionnaire ('I'): These are not affected by order of testing.

The analysis of variance was made with the following ideas in mind: (i) It provided a method, independent of the experimenters' biases, of classifying the data; (ii) the single individual was regarded as a source of a population of responses, and (iii) the residual was used as an estimate of error to provide a basis for comparing variances. We realized that both our 'significant' and 'non-significant' findings might eventually prove to be due to factors which were irrelevant to our purpose. We have, therefore, presented our raw data in Table 3, and have confined our attentions to those effects which have produced relatively large variances.

RESULTS AND DISCUSSION

Comparison of the treatments in relation to triad score.

While the triads appear to differ substantially among themselves (Table 4) they do not appear to be affected differently by the two types of treatments. None of the interactions of the triads in which the two treatments are involved are much larger

N = Non-directive
 T = Training
 I = 'Immediate' measures
 D = 'Delayed' measures

Average scores on the personal questionnaire
 on 17 occasions of testing

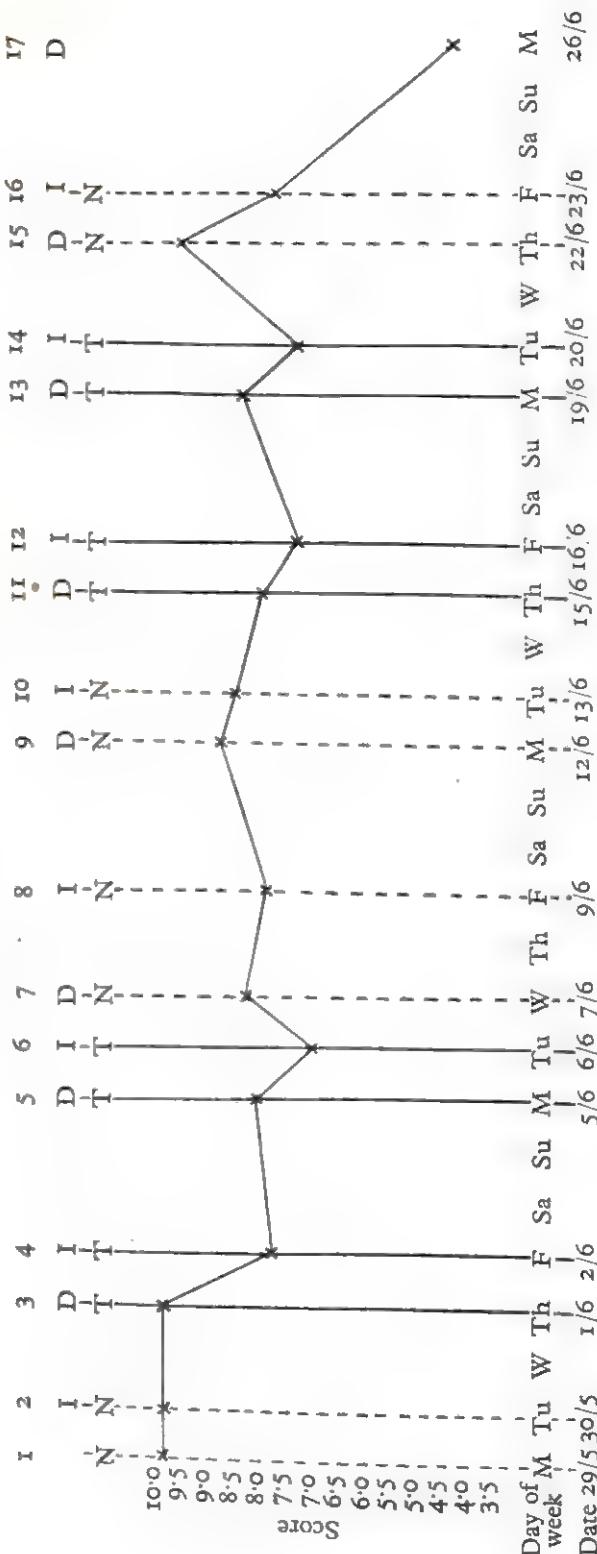


Table 3. Raw data

Triad No.	Occurrences														Summary of content		
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	X
1	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10.0
2	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10.0
3	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	9.913
4	10	10	7	10	10	10	10	10	10	10	7	10	10	10	10	10	9.63
5	10	10	10	10	10	10	10	10	10	10	10	7	10	7	7	7	9.25
6	10	10	10	10	10	10	10	10	10	10	7	10	5	10	10	2	9.0
\bar{x}	10	10	9.5	10	10	10	10	10	10	10	10	9.0	9.5	8.7	9.5	8.2	
7	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	5	9.6875
8	10	10	10	10	10	10	10	7	10	10	10	6	10	10	10	2	9.0625
\bar{x}	10	10	10	10	10	10	10	8.5	10	10	10	8.0	10	10	10	3.5	
9	10	10	7	10	10	7	10	10	7	10	7	10	7	10	7	5	8.5625
10	10	10	7	7	10	10	10	10	10	7	10	10	10	6	10	2	8.5
11	10	10	10	10	5	10	7	10	10	7	10	7	10	7	10	2	8.25
12	10	10	7	7	10	7	10	7	5	7	5	10	7	10	7	2	7.375
13	10	10	7	7	7	7	7	7	7	7	7	7	10	7	10	2	7.25
14	10	10	7	7	7	7	7	7	7	7	7	7	7	10	7	2	7.125
15	10	10	5	7	5	5	7	7	10	7	7	7	7	10	5	2	6.9375
16	10	10	7	7	5	7	5	10	7	5	7	7	5	10	7	2	6.9375
17	10	10	7	7	6	10	7	7	5	7	7	5	7	7	7	2	6.9375
18	10	10	7	7	2	7	10	10	7	7	5	7	5	7	7	2	6.875
19	10	7	7	5	7	7	7	7	7	7	5	7	7	7	7	2	6.625
\bar{x}	10	9.7	7.4	7.3	6.0	8.2	7.4	8.4	7.9	7.2	6.6	7.6	6.5	9.2	6.7	2.3	
20	10	10	7	5	5	7	5	5	7	5	5	5	10	5	10	2	6.4375
21	10	10	7	10	5	7	7	10	5	5	5	2	10	2	2	2	6.375
22	7	10	5	5	2	5	7	10	7	5	5	7	7	7	7	2	6.0
23	10	10	2	7	2	2	5	2	7	5	5	2	10	5	2	2	5.0625
\bar{x}	9.2	10	5.2	6.7	3.5	5.2	6.0	6.7	6.5	5.0	5.0	6.2	4.0	9.2	4.7	2.0	

Each column refers to an occasion of testing and each numbered row to a triad. Thus the raw score for the 10th triad on the 16th occasion of testing was 6. The triads have been divided into four classes, in accordance with the discussion on p. 89. The five rows marked \bar{x} give the means for each of these classes for each occasion, and the column marked X gives the mean for each triad over the 16 occasions.

than the residual. (See the interactions in Table 4 involving both 'K' and 'I'.) It was decided, therefore, to compare the two treatments in terms of the mean scores of all twenty-three items: i.e. to regard each triad as an imperfect measure of an entity common to all items, the entity being the illness as a whole.

Comparison of the triads with each-other, independently of treatment

The main effect for the triads produced a variance of 34.061, which is quite substantial and highly significant when compared with the residual (Table 4).

Table 4. *Analysis of variance*

No.	Source	DF	Var.	F. ratio
1	I	22	34.06	15.11§
2	K	1	1.70	
3	L	1	4.13	1.83
4	H	1	105.46	46.79§
5	W	1	105.46	46.79§
6	IK	22	2.24*	
7	IL	22	1.23*	
8	IH	22	3.41*	
9	IW	22	2.12*	
10	KL	1	103.33	45.84§
11	KH	1	120.98	53.67§
12	KW	1	83.22	36.92§
13	LH	1	14.48	6.42†
14	LW	1	29.96	13.29
15	HW	1	17.83*	
16	IKL	22	2.25*	
17	IKH	22	3.84*	
18	IKW	22	2.99*	
19	ILH	22	1.31*	
20	ILW	22	1.31*	
21	IHW	22	1.62*	
22	KLH	1	61.96	27.49§
23	KLW	1	.07*	
24	KHW	1	.00*	
25	LHW	1	23.50	10.43
26	IKLH	22	2.45*	
27	IKLW	22	1.47*	
28	IKHW	22	2.25*	
29	ILHW	22	1.42*	
30	KLHW	1	15.29	6.78†
31	IKLHW	22	2.82*	
	(original residual)			
32	Total	367		
33	New residual	334	2.25	

I = Triads; K = Kinds of treatment; L = Immediate and delayed; H = First and second half; W = Mid-week and weekend.

* Included in new residual, which includes all non-significant and experimentally irrelevant variances.

$$p = 0.05 \quad 0.025 \quad 0.01 \quad 0.001$$

Based on new residual.

This finding seems to have two implications. The first is that the rate of drop in score must have varied considerably from triad to triad. On the second and third occasion, a maximum score of 10 was provided by twenty-two out of the twenty-three triads. By the seventeenth occasion some items still had a score of 10, while others had dropped to the minimum score of 2. The second implication is that the triads with the highest mean score showing least changes and the triads with the lowest mean scores showing most change are classifiable in terms of their obvious psychological content. Six of the eight triads with a mean score of 9.0 or more (Nos. 1, 2, 3, 4, 5, 6, in Table 3) could be regarded as phobic in content.

The two remaining triads seem to involve parental concern. All triads with a mean score of 6.5 or less (Nos. 20, 21, 22 and 23) are depressive in content (see Table 2). All the remaining triads, except numbers nine and fourteen, could be called anxiety symptoms. This triple classification might have interacted with the two kinds of treatment if it had been introduced into the analysis. The relatively independent fluctuation of phobic and depressive symptoms has been reported in two other patients (Shapiro *et al.*, 1962; Shapiro, 1962).

Comparison of the two kinds of treatment

As was stated above, it was decided to consider the effects of the two kinds of treatment in terms of the mean scores of the twenty-three items. There is no overall difference between the two kinds of treatment in terms of mean scores. The mean variance for this main effect is only 1.698 ('K' Row 2, Table 4) which is less than the residual of 2.254. The mean for psychiatrist was 7.832 and for psychologists 7.967. However, substantial differences between the two kinds of treatment, warranting further attention, emerge when we take into account other important variables. 'K' has relatively large and significant interactions with 'L' (Row 10), with 'H' (Row 11) and with 'W' (Row 12), the variances being 103.3, 121.0 and 83.2 respectively.

Relation to immediate and delayed measurements (KL)

The mean for the 'immediate' effects of 'T' treatment is, on the average, smaller than the mean for the immediate effects of the 'N' treatment (Table 5), and the opposite is the case for the 'delayed' effects (Table 5). The implications of this finding are seen in the graph. In the case of the psychologists' treatments, each of the four 'immediate' measures (Nos. 4, 6, 12 and 14 in the Graph) is lower than its preceding measure; i.e. the immediate effect of each session of 'T' treatment appears to be that of improvement, the average improvement being 1.2715. On the other hand, every 'delayed' measure following an 'immediate' measure shows a relapse, the average deterioration being 1.2713. The 'T' treatments appear to have made no contribution to the *S* apparent substantial overall improvement, from extreme illness to an approach to recovery.

The essential contribution to this improvement must come from the 'N' treatment. Like the 'T' treatment, each of the four 'immediate' measures of the effects of 'N' treatments (Nos. 2, 8, 10 and 16), with the exception of the first measure of no change, showed improvement. The mean improvement was 0.652. The

Table 5. *Differences between treatments in relation to delay (KL) to part of experiment (KH) and part of week (KW)*

(Average rank order of testing shown in brackets)

KL	KH		KW	
	First half	Second half	Weekend	Mid-week
Immediate (9.0) 8.391	Delay (10.0) 7.543	(5.5) 9.076	(13.5) 6.859	(6.5) 8.978
'Non-Directive' means (9.0) 7.196	(10.0) 8.467	(5.5) 7.793	(8.5) 7.870	(10.5) 7.891
'Rational-Training' means N - T	-0.924	1.283	-1.011	-0.815
				1.087

A difference of 0.8394 is significant at the 0.001 level and of 0.6241 at the 0.01 level when compared with the 'new residual' of 2.25.

'delayed' measures of the 'N' treatment (Nos. 3, 9, 11 and 17) produced different and varying results, the first measure showing no change, the second a small relapse, the third a small improvement, and last a very large, and, in fact, the largest recorded change in the experiment. The mean improvement was 0.847, a finding directly opposite to that for the 'T' treatment.

Discussion of 'KL'

Our finding concerning the immediate effects of treatment sessions is quite different from that reported by Shapiro and Ravenette (1959) concerning paranoid beliefs. They found no immediate effects. The questionnaire results for the present subject were consistent with our subjective impressions. She often met the experimenters feeling and looking very ill, and by the end of the session seemed to be very much better. She spontaneously remarked on the process, and reported that another activity, typing, also made her feel temporarily better. The different results in the two experiments might stem from the fact that in the present experiment we dealt with feelings and in the other, beliefs. It is known clinically that some patients can be temporarily distracted from their depressions.

A second point is that we have repeated the finding of Shapiro and Ravenette that the relatively long term effects of a session of psychological treatment were different from the immediate effects, and that the direction of the long term effects is a function of the content of the immediately preceding treatment session. The difference in the present case might be due to 'T' treatment only having had a temporary distracting effect upon the *S*. It is, of course, conceivable that the introduction of 'T' treatment facilitated the effects of the psychiatrist's treatment. The fact remains that, in this situation, the patient reacted differently to the two kinds of treatment. *S* spontaneously indicated, towards the end of the experiment, that she depended upon the 'N' treatment.

It is difficult to assess the exact importance of the fact that different persons were involved in the two kinds of treatment, as *S* was in the habit of proclaiming both her considerable regard for the psychologists, and her hostility for I.M.M.

Relation to position during the week ('KW')

The means for the 'N' treatment were significantly lower than those for the 'T' treatment on weekend measures, and the opposite was the case for mid-week measures (Table 5).

This interaction is difficult to interpret, because the direction of difference between the two pairs of means is exactly the same as the difference in direction of their average rank orders of testing (Table 5). There seems no point, therefore, in speculating upon this result.

The relation to first and second halves of the experiment ('KH')

Measures of the effects of 'T' treatment were, in the first half, significantly lower than those for the 'N' treatment; this order was reversed in the second half. This could have been produced by (i) M.B.S.'s withdrawal from the second half of

the experiment; and (ii) by the change in character of the 'T' treatment after the first four sessions (see above description of treatment).

Effect of different testers

The mean score of all the eight tests administered by I.M.M. was 8.817 and, for the eight tests administered by Mr Mordal, was 7.832. A separate analysis of variance was carried out on the data provided by occasions 1 to 16 inclusive, because this set of sessions provided a balanced design for the two testers. The results (Table 6) indicate that the differences between the two testers might be reliable, the ratio of the 'between testers' variance to the residual variance being 37.263 (a/d , Table 6). However, the 'between tester' variance might not be significantly greater than the 'within testers' variance. The ratio is 3.621 (a/b , Table 5) which on 1 and 14 degrees of freedom does not reach the 0.05 level.

Table 6. *Differences between tests. Analysis of variance*

	d ₅	Var.
a Between testers	1	72.198
b Between occasions within testers	14	19.9409
c Between items	22	28.2095
d Residual	330	1.9375

In any case it would seem that the observed difference between the two testers could not have affected our main conclusions concerning differences between the delayed effects of the two kinds of treatment. The four measures of the delayed effects of each treatment were made up of both kinds of tests: two each in the case of 'rational-training', and three by Mordal and one by I.M.M. in the case of 'non-directive' treatment. Only one out of four pairs of measures was, therefore, made by different testers. Taking the observed overall mean difference of 0.985 between the two testers on sessions 1 to 16 as our basis, we can arrive at an expected difference of 0.246 between the delayed effects of the two treatments (0.8855/4). This is much less than the observed difference of 0.924 (see Table 5).

Additional considerations

1. The overall clinical improvement shown by the questionnaire results is consistent with our subjective impressions of the S's clinical condition. This improvement, with occasional relapses, continued for another four months.
2. S menstruated during two parts of the experiment, from 29/5 to 1/6 inclusive, and from 21/6 to 27/6 inclusive. The four highest scores in the experiment were produced during menstruation, and so were two of the lowest scores. S thought she was at her worst just before her periods. These data do not appear to affect our conclusions.
3. The course of Imipramine might have facilitated S's reactions to the two kinds of treatment.

CONCLUSION

The outstanding defect of the present experiment is that it was not continued for a longer period, possibly with systematic variation of independent variables derived

from the findings of the initial experiment. Such a long term experiment has just been completed on another patient (Shapiro *et al.*, 1962). Circumstances made this impossible in the present case. However, the present experiment does appear to have some value as a pilot for a certain method of research, and does appear to demonstrate that the systematic therapeutic manipulation of complex psychological variables can bring about meaningful changes in the psychological complaints of an individual patient.

Mrs P. Levine and Mrs R. Brook prepared the script. Miss N. Hemsley carried out a major part of the computations. Dr F. Post provided facilities, encouragement and criticism. A large part of the research was made possible by a grant from the Medical Research Council.

REFERENCES

JONES, H. G. (1956). Application of conditioning and learning techniques to the treatment of a psychiatric patient. *J. abnorm soc. Psychol.*, **52**, 414-419.

MAXWELL, A. E. (1958). *Experimental Design in Psychology and the Medical Sciences*. Methuen & Co.

PHILLIPS, J. P. N. (1961). Technical Aspects of the Questionnaire (unpublished).

SHAPIRO, M. B. (1961a). A method of measuring psychological changes specific to the individual psychiatric patient. *Brit. J. med. Psychol.*, **34**, 151-155.

SHAPIRO, M. B. (1961b). The Personal Questionnaire Abbreviated Manual. Unpubl.

SHAPIRO, M. B. (1962). Clinical approach to fundamental research with special reference to the single case. In *Basic Research Techniques in Psychiatry* (eds. P. Sainsbury, & N. Kreitman). Oxford University Press. (In Press.)

SHAPIRO, M. B. & RAVENETTE, A. T. (1959). A preliminary experiment of paranoid delusions. *J. ment. Sci.*, **105**, 295-312.

SHAPIRO, M. B., NEUFELD, I. & POST, F. (1962). Note: Experimental study of depressive illness. *Psychological Reports*, **10**, 590.

WOLPE, J. (1958). *Psychotherapy by Reciprocal Inhibition*. Stanford University Press.

Manuscript received 31 May 1962

Changes Due to Noise in the Threshold of Fusion of Paired Light Flashes in Schizophrenics and Normals

By P. H. VENABLES

*Medical Research Council, Social Psychiatry Research Unit, Institute of Psychiatry,
London, S.E.5*

The threshold inter-flash interval at which pairs of flashes appear to fuse to produce the appearance of a single flash was measured in schizophrenic and normal subjects under quiet and noise conditions. The results show that while the performance of normal subjects is not affected by noise that of schizophrenics is. The amount of change of threshold due to noise shown by the schizophrenics is closely related to their threshold in quiet. The results are discussed in terms of the influence of a regulatory system on the cortical activity of normal subjects which is lacking in schizophrenics.

1. INTRODUCTION

In a previous paper (Venables, 1963) results were presented which tended to show that measures of autonomic and cortical activity levels were negatively related in a group of normal subjects and positively related in a group of schizophrenics. Thus those schizophrenic subjects who showed high cortical activity (measured by their ability to see as two, pairs of light flashes which were temporally close) also showed high autonomic activity, while in normals those who showed high cortical activity tended to show low autonomic activity.

A tentative explanatory hypothesis for these results suggested that in normal subjects there was an inhibitory mechanism operating between cortex and lower centres and that this was functionally under-active in schizophrenics. That the direction of this inhibitory mechanism was from lower centres to the cortex was suggested by a preliminary finding that in normals noise appeared to disturb autonomic activity levels but had no effect on an index of cortical activity.

The aim of this study was to examine further the effects of noise on cortical activity. The effects of noise on autonomic activity are considered in another paper.

A dilemma presented by the findings cited above is that they are derived from two types of experiment. In the first subjects are compared and in the second the same subject is studied on two occasions.

The first, static, picture of a negative relationship between cortical and autonomic activity levels presented by a group of normal subjects suggests the presence of an inhibitory link between cortex and lower centres.

On the other hand the second, dynamic, set of results suggest rather a regulatory influence on the cortex which tends to maintain its level of function constant in normal subjects.

The results are therefore examined with the aim of discovering whether they are

more consistent with the idea of an inhibitory or a regulatory mechanism in normal subjects which is absent in schizophrenics.

The measure which is used in this study as an index of cortical activity is the ability of the cortex to resolve pairs of light flashes which are temporally close. Lindsley (1957) was able to show that by raising the general level of activity of the optic cortex by direct reticular stimulation he was able to achieve resolution into two spikes of a single spike of evoked potential resulting from a pair of light flashes 50 msec apart. Similarly Jung (1957) was able to show that by raising the activity level of individual neurons by repeated thalamic and reticular stimuli causing some non-specific neuron discharges, the C.F.F. of these neurons may be increased to 10-15 sec higher than under control conditions.

The task presented to the subject was that of reporting his subjective threshold of fusion of pairs of light flashes. A threshold inter-flash interval was obtained for each subject in quiet conditions and later after exposure to continuous white noise.

2. METHOD

Apparatus

Approximately square wave flashes of light 5 msec in duration were produced by applying pulses to the grid of a Ferranti CL.66 cathode ray tube. This tube has a phosphor which produces an essentially white source of light and has a decay time of 5 μ sec. Associated timing equipment made possible the presentation of two flashes of light, the interval between which was continuously variable and could be set to the nearest millisecond in the working range of the experiment, between 30 and 140 msec. The intensity of the flash was attenuated to 50 ft L by a neutral density filter. An edge lit perspex block in front of the tube face provided a fixation patch of 0.025 ft L .

A mask with a circular hole in front of the tube face gave a visual angle of 1.8°. The experimental room was completely darkened and the timing apparatus was operated from an adjacent room. A simple press button signalling system enabled the subject to report the number of flashes he saw.

Procedure

The subject was seated with his eyes approximately 4 ft from the tube face, he was asked to rest his head between two pads on the chair back which kept it in line with the optical axis of the apparatus. He was told that he would see either one or two flashes of light and that all he had to do was to report the number by giving one or two presses on a button conveniently situated at his right hand. A preliminary practice was given until the subject felt confident that he could do what was required.

The threshold of fusion was determined by the following technique. A pair of flashes with a long inter-flash interval (I.F.I.) was presented; if the subject pressed twice, flashes with a short I.F.I. were presented; if the subject then pressed once the procedure was repeated with decreasing range until two I.F.I.s a millisecond apart were reached. Flashes at these I.F.I.s were then twice repeated and if the responses were maintained the I.F.I. at which the subject pressed once was taken as the threshold.

The following sequence is from a typical record:

100-2, 50-1, 90-2, 60-1, 70-1, 80-2, 70-1, 75-1, 78-2, 75-1, 76-1, 76-2, 75-1, 75-1, 76-2, 75-1, 76-2.

Experience with the technique showed that it was important to obtain the threshold with the minimum number of trials. The subject tended to make more random responses after about twenty presentations of the paired flashes. With care few subjects found difficulty in giving a threshold to the nearest millisecond.

Other work (for example Venables, 1963), tended to show that thresholds over 140 msec were unreliable. The data from five schizophrenic and two normal subjects with thresholds in excess of this value have been discarded.

The subject sat for four minutes in the dark prior to the determination of the first threshold. This period was to allow for adaptation to the intensity of the fixation patch. Following this a minute elapsed before an 80 db white noise was presented from a loudspeaker 4 ft in front of the subject and below the flash source. Five seconds were allowed to increase the noise to its full intensity which was then maintained for three minutes. A second threshold was then taken at the end of the period while the noise was still heard.

Subjects

Sixty-three chronic schizophrenic subjects gave satisfactory thresholds in quiet and noise conditions. Of these 38 were receiving no medication whatsoever, and 25 were on various dosages of chlorpromazine. Seven of the unmedicated patients were women, all those on chlorpromazine were men. No patients were tested who fell in the category of those showing coherently expressed paranoid delusions who had previously (Venables, 1963) been shown to differ from the remainder of schizophrenics. Forty-seven normal subjects were tested; 28 were men and 19 women. No differences due to sex could be found in any of the data. All subjects were aged between 20 and 60 years.

3. RESULTS

The results are summarized in Table 1. The upper half of the table compares the results from those patients who were, or were not receiving chlorpromazine. As these groups were not found to differ on any of the variables presented their data were combined and are shown in the lower half of the table where they are compared with the results from the normal group.

A simple comparison of mean initial threshold and mean changes in threshold due to noise showed no difference between schizophrenics and normals (Cols. 3 and 4).

Plots of the data relating initial threshold in quiet and change in threshold due to noise were made for schizophrenics and normals. The amount of change shown by normals varied randomly about zero. There was no relationship between threshold change and initial threshold (Cols. 5, 6 & 7). In the schizophrenics, however, the amount of change of threshold brought about by exposure to noise was closely ($p < 0.001$) related to initial threshold in quiet. Those patients who started with a low threshold tended to show an increased threshold in noise while those who started with a high threshold exhibited a lowered threshold in noise. Using the terms of the regression line it was shown that zero change in threshold was estimated to be given by those subjects having an initial threshold of 65 msec. Those with thresholds above this value tended to show a decreased threshold in noise while those with a threshold less than 65 msec showed a raised threshold in noise.

The final column of the table shows the standard error of the change of threshold (y) calculated from the initial threshold (x). The size of this error is significantly larger ($p < 0.02$) in the schizophrenic data than it is in the normal.

4. DISCUSSION

The results show that noise has no significant effect on the two-flash threshold of normal subjects. This would suggest that in so far as this measure is an index of cortical activity, either some regulatory function maintains the level of cortical

Table I. Summary of data relating initial threshold and change in threshold due to noise in schizophrenics and normals

	1	2	3	4	5	6	7	8
Groups					Correlation between threshold and change in threshold		Regression line relating change in threshold (y) to initial threshold (x)	
Schizophrenics not on drugs	N	Mean (msec.) initial threshold	Mean change in threshold msec.		$y = bx + a$		$\sigma_{y,x}$	
Significance of difference	38	70.6 S.D. 20.8 $p > 0.05$	-0.7 S.D. 10.4 $p > 0.05$	-0.62	-0.31	+21.2	8.1	
Schizophrenics on chlorpromazine	25	65.0 S.D. 20.7	-1.4 S.D. 9.2	$p > 0.05$	$p > 0.05$	$p > 0.05$		
Schizophrenics combined groups	63	68.3 S.D. 20.9 $p > 0.05$	-0.9 S.D. 9.9 $p > 0.05$	-0.63	-0.30	+19.5	6.8	
Significance of difference	47	71.5 S.D. 16.7	+0.02 S.D. 5.7	-0.08	-0.03	+4.8	5.7	
Normals								

activity at some relatively constant value or that the threshold at which noise has an effect if at all is higher than 80 db.

In chronic schizophrenics it would appear that the point at which noise has an effect is reached at 80 db. At this point there appears to be an absence in schizophrenics of the regulatory function seen in normals.

An emphasis is placed on a regulatory rather than inhibitory function in normals because it would be necessary to produce zero change in threshold in schizophrenics to have an excitatory mechanism operating in those patients with thresholds below 65 msec and an inhibitory process in those with higher initial thresholds.

The hypothesis which suggests that it is lack of regulatory control which is exemplified by the schizophrenic is only one of a number which are possible.

It must be noted that in the schizophrenic we do not have a mechanism which is unstable. The effect of noise is to change the thresholds of individual patients towards a central value of 65 msec. It is possible, therefore, that there are perhaps two stages of effective control on level of cortical activity. A final stage operative in the normal maintains the cortex at a relatively fixed value of activity. This may be relatively non-operative in the schizophrenic. In the schizophrenic we may be seeing what is effectively a second stage of control whose function is to prevent ultimate instability of cortical activity.

On the other hand the results given by the schizophrenic may exemplify the malfunctioning of a regulatory system rather than its absence. We might postulate, following a Pavlovian type of hypothesis, a tendency in the schizophrenic for stimulation to produce a relative excess of excitation which in turn leads to the development of protective inhibition. Thus at low levels of cortical activity we have the tendency to an increase in activity. With higher levels this tendency to further excitation brings about the onset of protective inhibition. At an activity level represented by a two-flash threshold of 65 msec, inhibition and excitation are balanced and with higher pre-stimulus activity levels inhibition is dominant and stimulation brings about a decrease in activity.

It should be noted that this latter suggestion fits quite well with the Pavlovian notion of the 'paradoxical effect' which has been shown in chronic schizophrenics in the visual but not the auditory modality (Venables & Tizard, 1956, 1958). A possible test of the consistency of the latter explanation with the 'paradoxical effect' findings would be to determine whether or not the fusion of auditory clicks would be influenced in schizophrenics by additional stimulation in another modality.

It must be added as a final caution that the results presented may apply only to particular intensities and timings of noise and flash and that as we are apparently dealing with a dynamic regulatory system the findings will probably not apply to other conditions without modification.

REFERENCES

JUNG, R. (1957). Chap. 21 in *Reticular Formation of the Brain* (eds. H. H. Jasper *et al.*). London: Churchill.

LINDSLEY, D. B. (1957). Chap. 25 in *Reticular Formation of the Brain* (eds. H. H. Jasper *et al.*). London: Churchill.

VENABLES, P. H. (1963). The relationship between level of skin potential and fusion of paired light flashes in schizophrenic and normal subjects. (To be published.)

VENABLES, P. H. & TIZARD, J. (1956). Paradoxical effects in the reaction time of schizophrenics. *J. abnorm. soc. Psychol.* **53**, 220-224.

VENABLES, P. H. & TIZARD, J. (1958). The effect of auditory stimulus intensity on the reaction time of schizophrenics. *J. ment. Sci.* **104**, 1160-1164.

Manuscript received 27 August 1962

Conceptual Thinking, Personality and Conditioning

By S. H. LOVIBOND

University of Adelaide

The GSRs of Ss with high and low scores on the Object Sorting Test, and the Extraversion and Neuroticism scales of the MPI are conditioned in both aversive and non-aversive situations. It is hypothesized that high scores on the Sorting Test result from a weak attention response and a concomitant weakness of the inhibitory process. In accordance with hypothesis high and low scorers on the Sorting Test are differentiated on a number of non-aversive conditioning variables. No consistent differences are found in aversive conditioning. Neither E nor N scores are significantly related to conditioning variables.

During recent years there has been a renewal of interest in the possibility of establishing links between conditioning and personality variables. The present interest in this area has resulted chiefly from the development of Eysenck's dimensional theory of personality and from Spence's extension of Hullian theory to include variables of importance in personality study.

In Eysenck's theory the dimension linking personality and conditioning is extraversion. The personality variable which Spence emphasizes in relation to differences in conditionability is anxiety or neuroticism.

In the experiment to be reported, each of these personality variables, extraversion and neuroticism, is examined in relation to conditioning measures. However, main interest centres on the relationship between conditioning measures and looseness of thinking as measured by the Object Sorting Test.

From previous work (Lovibond, 1953, 1954) it was concluded that the characteristic feature of schizophrenic thought disorder is a failure to inhibit associations irrelevant to the context. Accordingly a method was developed to permit analysis of Sorting Test responses in terms of the degree of inessentiality of the definitions offered. Each inessential response was scored 1-3, giving a total score range of 0 (indicating no impairment) to 57 (indicating severe impairment).

Subsequent unpublished studies have demonstrated both the reliability of the Sorting Test and the comparative stability of test scores under such influences as ECT and LSD.

In the original investigation a cutting score of 6 was found to differentiate 90 per cent of the controls and 60 per cent of the schizophrenics.

At that time the 10 per cent of high-scoring normals was believed to comprise preschizophrenics or ambulatory schizophrenics. However, recent testing has suggested that the proportion of high-scoring normals is considerably higher than 10 per cent. It is now hypothesized that the looseness of thinking measured by the Sorting Test exists in many members of the population who are not, have not been, and will not become, schizophrenic.

It follows that certain aspects of schizophrenic-type thinking may most conveniently be studied in those normals who obtain high scores on the Sorting Test, but whose thinking is free from the complications of mental disorder.

In attempting to understand the processes giving rise to high scores on the Sorting Test, Pavlovian concepts of higher nervous activity have been used as a source of hypotheses.

In the original formulation it was assumed that the presentation of the verbal instructions of the Sorting Test, together with the total context of the test, initiates a process of excitation which activates linkages in the subjects' conceptual structures relevant to the solution of the task. At the same time linkages irrelevant to the performance of the task are inhibited by a process of negative induction.

It is now further assumed that the negative induction of irrelevant linkages can be achieved only on the basis of a relatively strong process of excitation.

In this connection Kaada and Johannessen (1960), following earlier Soviet workers, have recently drawn attention to the inhibitory effects which accompany the 'arousal response' and which may be elicited from stimulation of widespread areas of the cortex. These workers point out that in addition to excitatory effects, adequate novel stimulation produces in the intact animal an inhibition of ongoing activity and a directing of attention to the source of stimulation.

For these reasons the term 'behaviour attention response' is preferred and it is regarded as being in all probability identical with Pavlov's 'orienting reflex'.

In these terms high scores on the Sorting Test result from a weak attention response, or a failure to inhibit irrelevant linkages as a concomitant of a low level of cortical excitability. This low level of cortical excitability may or may not be accompanied by or result from a low level of subcortical excitability.

On the basis of the foregoing analysis it is possible to derive a number of testable hypotheses concerning the differential performance of high and low scorers on the Sorting Test in a conditioning situation. In particular, high scorers should show lower reactivity including poorer conditioning, and they should also differentiate less readily than low-scoring subjects.

A pilot experiment carried out with 17 high Sorting-Test scorers and 16 low Sorting-Test scorers drawn from Psychology I students of the University of Adelaide confirmed these predictions. The following set of interrelated hypotheses was then formulated for testing in the main study.

In a GSR conditioning situation high scorers on the Sorting Test will show:

- (1) a lower initial reaction to the neutral stimuli during the adaptation period;
- (2) a slower rate of adaptation to the neutral stimuli;
- (3) less disinhibition of adaptation following the first reinforcement by the unconditioned stimulus;
- (4) a slower rate of conditioning;
- (5) a smaller degree of inhibition with reinforcement;
- (6) a slower rate of differentiation;
- (7) less reminiscence or disinhibition of inhibition with reinforcement following a change in the stimulus conditions;
- (8) less ability to reverse the responses to the conditioned stimulus and the differentiated stimulus when the differentiated stimulus is reinforced in place of the conditioned stimulus.

It was further predicted that these hypothesized effects would be more marked with non-aversive than with aversive conditioning on the grounds that aversive stimulation, particularly electric shock, is likely to mask differences between groups as a result of strong augmentation of unconditioned responses to the neutral stimuli (Lovibond 1961).

Extraversion

According to Eysenck, extraverts develop excitatory potentials slowly, and develop inhibitory potentials quickly. Hence extraverts should condition poorly.

A number of tests of this prediction have been made, chiefly with eye-blink conditioning. In an early study, Franks (1957) obtained a negative correlation of -0.46 between extraversion and conditioning. Later workers have not found such a strong relationship, the average correlation from half a dozen studies (Willett, 1950, Becker, 1960) being about -0.2 .

Inasmuch as Eysenck's hypothesis assumes a general factor of conditionability, the poorer conditioning performance of extraverts should be observed with both aversive and non-aversive procedures.

Neuroticism

Leading theorists have offered conflicting predictions concerning the relationship between neuroticism and conditioning. The Iowa group (Spence, 1953; Spence & Farber, 1953; Spence & Taylor, 1951) has assumed that neuroticism, emotionality or anxiety acts as a drive which summates with other drives in the activation of habit strength to produce performance. Hence, in simple conditioning situations, both aversive and non-aversive, high N subjects should condition better than those low on N.

On the other hand, Hilgard, Jones & Kaplan (1951) have proposed that anxiety or neuroticism acts as a drive which is specific to threatening situations.

Hence, high N subjects should condition better than low N subjects with aversive stimuli, but there should be no difference between these two groups in the case of non-aversive conditioning.

The prediction of Eysenck is that there should be no difference between high and low N subjects in either aversive or non-aversive conditioning.

Since the dimension relating to conditionability is introversion-extraversion, and this dimension is theoretically independent of neuroticism, there should be no relationship between an adequate measure of neuroticism and any measure of conditionability.

No adequate test of these hypotheses has yet been reported in the literature. Bindra, Paterson & Strzelecki (1955) have reported a study in which an attempt was made to decide between the Spence-Taylor-Farber and Hilgard-Jones-Kaplan hypotheses by comparing the performance of high and low anxious subjects in a non-aversive conditioning situation, namely salivary conditioning. However, a careful study by Willett (1960) has thrown considerable doubt on the validity of conventional measures of salivary conditioning in human beings, and a replication

of the Bindra study, with a control group, in the Adelaide laboratory failed to demonstrate a difference in the performance of experimental subjects and controls.

The experiment to be reported was designed to test each of the foregoing hypotheses relating scores on the Sorting Test, and extraversion and neuroticism to conditioning. Two GSR conditioning situations were used. In one conditioning session an electric shock to the fingers served as the UCS, and in the other session the GSR was elicited unconditionally in the same subjects by a series of coloured slides of nude females.

METHOD

Subjects

The initial pool of subjects for the investigation comprised the youngest 100 male students taking Psychology Part I in the University of Adelaide during 1961. In order to obtain 100 subjects it was necessary to extend the upper age limit to 26 years. The 100 subjects were given the 24 item forms of the Maudsley Personality Inventory E and N scales and Part II of the Object Sorting Test. Part II of the Sorting Test had previously been shown to predict external criteria almost as well as the full scale and it could be administered in about one third of the time required for the full scale. In the pilot experiment the full scale of the Sorting Test was used. In both experiments the Sorting Test was administered in almost all cases by senior students.

Table 1. Score ranges and median scores of the high, medium and low groups on the three tests

Test	Group	N	Score Range	Median
Sorting Test	High	20	5-12	6
	Medium	25	1-4	2
	Low	20	0	0
E Scale	High	20	34-45	38.5
	Medium	25	20-33	27
	Low	20	5-19	15
N Scale	High	20	31-44	35
	Medium	25	15-30	21
	Low	20	4-14	8

Completed test records were available from 93 subjects. From this group the 20 highest scorers and the 20 lowest scorers on each of the three tests, Extraversion, Neuroticism and Sorting Test were selected. Overlapping reduced the total number of subjects to 65. Thus on each of the three test variables there were available 20 subjects with high scores, 20 with low scores and 25 with intermediate scores.

Table 2. Intercorrelations between the three test variables

Test	(N = 65)		
	ST	E	N
Sorting Test	—	-0.33	-0.09
Extraversion		—	-0.26
Neuroticism			—

The mean scores and standard deviations of the original 93 subjects on the E and N scales were almost identical with those of the American students in the standardization group. Scores on the Sorting Test showed marked positive skewing. The median scores and score ranges

of the high, medium and low groups on the three tests are shown in Table 1 and the inter-correlations between the tests are presented in Table 2.

Table 2 reveals that as in the standardization group there is a negative correlation between E and N. This correlation results from the tendency for extremely introverted Ss to obtain above average N scores.

Sorting Test and N scores are virtually independent, but there is an appreciable and significant negative correlation between ST and E. This latter result is in contradiction with that obtained from the smaller pilot sample, for whom ST and E were positively correlated ($r = -0.35$).

Apparatus

The circuit used to measure the GSR was that described by Davis, Siddons & Stout (1954) modified to drive a Both pen recorder. The degree of attenuation employed varied from S to S between 100 and 200 ohms per mm of pen deflection. The Ss palm was rubbed thoroughly with electrode jelly before attaching the electrodes which were placed as far apart as practicable on the palm of the left hand. The 50-cycle a.c. electric shock which served as one UCS was produced by a constant current electronic stimulator and was delivered to the tips of the second and fourth fingers.

The nude female slides were projected in stereoscopic pairs through polaroid filters by means of twin Paximat 35 mm projectors with automatic slide changers. When viewing the slides Ss wore spectacles with appropriate polaroid lenses. The stimuli to be conditioned and differentiated were tones and lights. In one session the CS was a 500 cycle tone and the DS was a 1500 cycle tone. These tones were generated by two audio oscillators and fed through earphones at approximately 60 db above threshold.

In the other session the CS was a red light and the DS was a white light. The light stimuli were provided by two 32 V, 15 W lamps mounted side by side in a box and arranged to illuminate two circular ground glass apertures 3 in. in diameter and 9 in. apart in the front panel of the box. The entire sequence of stimulus events was presented automatically by a system of electronic timers and switching apparatus.

The task of the experimenter, once the session was under way, was simply to monitor the GSR recorder and make any necessary adjustments to the attenuation and reset switches. Two adjacent rooms served as the experimental and controls rooms. All apparatus was housed in the control room and S sat in an armchair facing a 6 ft \times 6 ft metallic screen on which the slides were shown and in front of which the light box was placed. Projection was through holes cut in the one-way vision glass which separated the two rooms. The hum of the recorder and a ventilating fan served to mask slight sounds from the control room and outside sources.

Procedure

The method of delayed conditioning was used with an 11-second interval between CS onset and UCS onset. Delayed conditioning was used from the first reinforcement to provide a measure of the response to the CS and the DS on every trial. The duration of the shock was one second and CS and shock offset were simultaneous.

The duration of the nude pictures was 10 seconds and in the case of this conditioning procedure the CS and UCS overlapped by one second. The CS and DS were presented in strict alternation with an intertrial interval, measured from CS onset to CS onset, of 55 seconds.

Strict alternation rather than random alternation was used to facilitate the development of a 'dynamic stereotype' (Pavlov, 1927) which was later to be reversed.

In the preliminary adaptation period, the CS and the DS were each presented twice without reinforcement. The fifth stimulus was the first reinforced presentation of the CS and it was followed by 11 further CS-UCS combinations alternated with 12 unreinforced DS presentations.

During aversive conditioning the DS was presented alone, but during non-aversive conditioning the DS was followed by a slide of an outdoor scene. A different nude was presented on each occasion following the CS.

The strength of the shock was set at 1.5 mA for the first two shocks and it was then increased to 2 mA for the remaining 10 reinforced trials. Ss post experimental descriptions of the strength of the shock ranged from mild to rather severe. Following the final DS (stimulus No. 28) a rest of 2 minutes was given. After this time six further presentations of each stimulus were

made but with the stimuli reversed, i.e. the former CS was now unreinforced and the former DS was reinforced.

To summarize the procedure, stimuli 1-4 were adaptation trials, odd number trials from 5-27 (i.e. 12 in all) were reinforced presentations of the CS. Even numbered trials from 6-28 were unreinforced presentations of the DS. On odd number trials from 29-39 the CS was unreinforced and on even number trials from 30-40 the DS was reinforced.

The interval between the two conditioning sessions was 2-7 days and type of CS and order of conditioning were counterbalanced.

Scoring

Unconditioned Responses. In order to check the possibility of differences in group responsiveness to the UCS, two measures were computed for each *S*. These were the sum of the responses to the third and fourth and to the eleventh and twelfth presentations of the UCS.

When no CR or a small CR occurred, measurement of the UCR presented no difficulties. However, when a sizable CR occurred it tended to mask the UCR. Hence a UCR was defined as the maximum deflection of the pen which occurred between CS onset and a point 10 seconds after UCS onset. The resulting value was not a pure measure of responsiveness to the UCS, but it appeared to be the most realistic measure in the absence of isolated presentations of the UCS.

Conditioning measures. The criterion of a conditioned GSR was change in conductance during the period of operation of the CS. The unit of conductance change employed was the square root of micromhos $\times 100$.

For the purposes of analysis eight separate scores relating to the eight hypotheses were computed for each *S*. These were

IR	<i>Initial Response:</i> response to the initial neutral stimulus.
AD	<i>Adaptation:</i> the regression of responses R_1 to R_5 .
D_1	<i>Disinhibition 1:</i> the difference between R_5 and R_6 .
RC	<i>Rate of conditioning:</i> the regression of responses 5, 7, 9, 11, 13 and 15.
DIF	<i>Differentiation:</i> the difference between the regressions of responses 6, 8, 10, 12, 14 and 16, and responses 5, 7, 9, 11, 13 and 15.
RI	<i>Reinforcement Inhibition:</i> the regression of responses 17, 19, 21, 23, 25 and 27.
D_2	<i>Disinhibition 2:</i> $R_{29}-R_{27}$.
REV	<i>Reversal:</i> reversal of responses to CS and DS: (R_{25} and 27-R ₂₆ and 28)—(R_{38} and 40-R ₃₇ and 39). The distributions of all the above measures were re-scaled by simple arithmetic to give a score range of 0-15 whilst retaining the original form of the distribution. In all cases high scores indicated strong excitation or inhibition.

RESULTS

UCS Responsiveness

The differences in UCS responsiveness between the high, medium and low subgroups were very small for all three tests. None of the differences approached significance.

Excitation and Inhibition measures.

In terms of the experimental hypotheses it was expected that all seven of the excitation and inhibition measures would intercorrelate positively, with the four excitation measures and the three inhibition measures possibly forming clusters with higher intercorrelations.

Pearson product moment correlations were computed to test these predictions, despite the fact that the score distributions were closer to half-normal than normal for most of the variables. The justification for using r under these circumstances was provided by the work of Heath (1961).

Heath correlated large numbers of successive pairs of samples drawn from one normal and one half normal distribution and from two half normal distributions. Remarkably little deviation from expected frequencies of high values of r was observed, and it was concluded that departure from normality in one or both variables does not vitiate the use of r provided the relationship is reasonably linear.

Since the aversive conditioning curve reached its asymptote after the first reinforcement, it was necessary to use a different measure of rate of conditioning with this method.

The measure adopted was the difference between the response to CS on trials 5 and 7. (It was recognized that this measure would in all probability include a large component of augmentation of the UCR to the CS.)

For similar reasons the aversive differentiation measure was based on the differences of the regressions of the responses to trials 7, 9, 11, 13 and 15 and 6, 8, 10, 12 and 14.

The intercorrelations between the non-aversive excitation and inhibition measures are presented in Table 3.

Table 3. *Intercorrelations between non-aversive excitation and inhibition measures*

	(N = 65)						
	D ₁	RC	D ₂	AD	DIF	RI	REV
IR	+0.33	+0.15	+0.02	+0.60	+0.30	-0.12	+0.18
D ₁		+0.44	+0.27	+0.60	+0.54	+0.38	-0.09
RC			+0.02	+0.27	+0.80	+0.13	+0.14
D ₂				+0.17	+0.11	+0.44	+0.10
AD					+0.42	+0.20	-0.08
DIF						+0.13	+0.16
RI							-0.19

It can be seen from Table 3 that if the reversal measure is excluded 20 of the 21 intercorrelations are positive, although in some cases the relationship is slight.

The highest correlations are between excitation and inhibition measures, but it must be emphasized that these measures are not independent.

Thus, the differentiation measure embodies the rate of conditioning measure, the possible size of the D₁ score is limited by the amount of adaptation and so on.

However, the coherence of the non-aversive measures can be judged by a comparison with the correlations for aversive conditioning.

In aversive conditioning eleven of the intercorrelations were positive, and ten were negative.

In this case also the few relatively high inter-correlations were explicable in terms of lack of independence of the measures.

The measure of capacity to reverse the responses to the CS and the DS was independent of both aversive and non-aversive conditioning measures.

Relationships between test and conditioning measures

Sorting Test. The correlations between Sorting Test scores and the seven non-aversive excitation and inhibition measures are contained in Table 4. Six of these

correlations are negative in conformity with the experimental hypotheses, four being significant beyond the 5 per cent level using the appropriate one-tail test. The reversal measure is uncorrelated with Sorting Test performance. The differential non-aversive conditioning performance of high and low scorers on the Sorting Test is illustrated in Figure 1. The curves in Figure 1 show the initially low response of the high group, the slower adaptation in this group and the relative lack of disinhibition of adaptation following the first reinforcement. A slow rate of conditioning and relatively little inhibition with reinforcement gives the high group an almost straight learning curve.

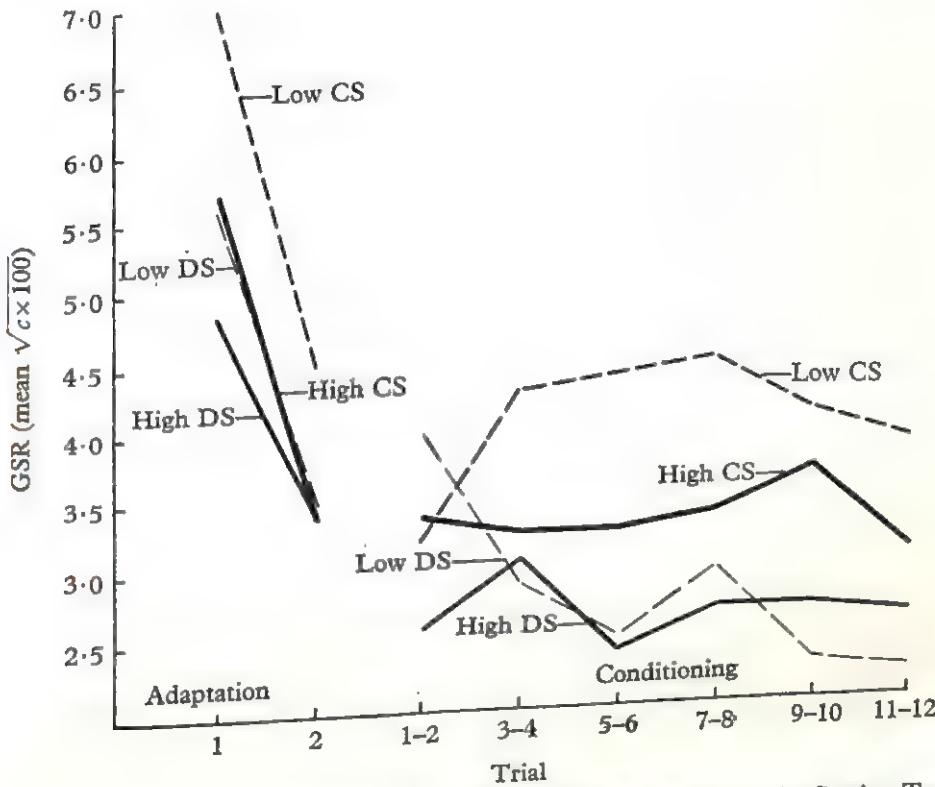


Fig. 1. Non-aversive conditioning curves of high and low scorers on the Sorting Test.

In the low group, by contrast, the CR increases rapidly to approach its maximum by the fourth trial followed by response decrement after the eighth trial.

The contrasting rates of differentiation for the two groups can also be seen clearly in Figure 1. The curves shown in Figure 1 are virtually identical in form to those obtained in the pilot experiment.

None of the correlations between Sorting Test scores and aversive conditioning measures was significant. Nevertheless, some features of the differential non-aversive conditioning performance of high and low scorers may be observed in the aversive conditioning curves shown in Figure 2. The slight differences during the adaptation period tend to be the reverse of those observed in non-aversive conditioning. However, following the first reinforcement greater disinhibition of adaptation is shown by

the low group. The rate of increase in response to the CS is also greater in the low group, which reaches its maximum level of response on the second trial. By contrast, the peak of the high group curve occurs on the fourth trial. In the case of the medium Sorting Test scorers (not shown) the maximum level of response to the CS occurred on the third trial.

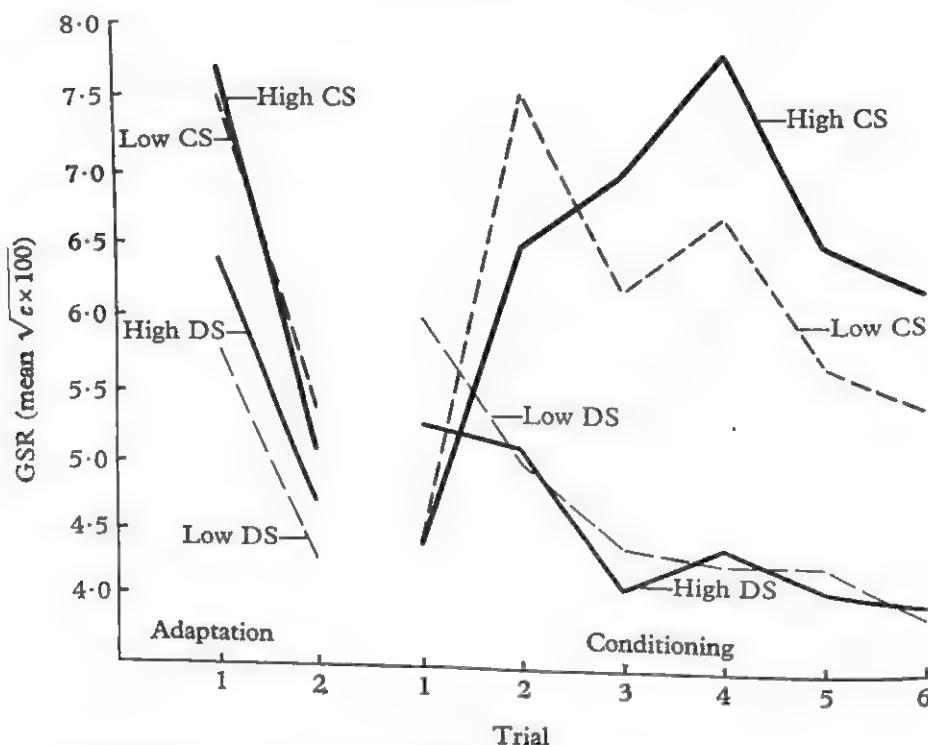


Fig. 2. Aversive conditioning curves of high and low scorers on the Sorting Test.

Extraversion. Table 4 shows no consistent tendency for scores on the E scale to correlate with non-aversive conditioning variables. Similarly none of the correlations between E scores and aversive conditioning measures was significant.

Table 4. *Correlations between test and non-aversive excitation and inhibition measures*

Test	(N = 65)							
	IR	D ₁	RC	D ₂	AD	DIF	RI	REV
Sorting Test	-0.21	-0.27	-0.22	+0.07	-0.15	-0.35	-0.02	+0.01
E Scale	+0.14	+0.02	-0.10	-0.04	+0.06	-0.10	-0.06	+0.06
N Scale	-0.04	-0.06	+0.10	-0.21	-0.11	+0.14	-0.15	-0.11

Neuroticism. The correlations between N and non-aversive conditioning variables in Table 4 are mainly negative. The positive relationships are with the important rate of conditioning and differentiation variables, but none of the correlations reaches the 5 per cent level of significance with a two-tail test.

In the case of aversive conditioning 6 of the 7 correlations between N and conditioning variables were negative and again all were insignificant. The only positive relationship was that with differentiation.

Conditioned time reflexes

With the invariant intertrial interval used in the experiment, there was some tendency for GSRs to occur in the period immediately prior to the onset of the CS and DS. Most of these responses were clearly conditioned reflexes to time which would be expected to disappear during the course of conditioning as a result of 'inhibition of delay'.

From the general theory relating Sorting Test scores and conditioning, it would be predicted that conditioned time reflexes would appear earlier and would disappear earlier in the case of the low-scoring group than in the case of the high-scoring group.

In order to test this prediction, a count was made of the number of responses made by each *S* during the 10 seconds before CS and DS onset in the first and last halves of the original conditioning period. In agreement with prediction, low scorers made more anticipatory responses than high scorers in the first half (low mean 1.5, high mean 1.3), and less in the second half (low mean 1.3 and high mean 1.7). Medium Sorting Test scorers occupied an intermediate position (means 1.6 and 1.6).

DISCUSSION

The outcome of the present study affords some encouragement for those who believe with Eysenck (1957) that the way forward in personality study is to seek integrations with molar neurophysiology and general behaviour theory and to develop rational rather than notional or empirical test procedures.

Some of the current formulations in this area have not been supported, but the specific hypotheses relating conditioning and performance on the Sorting Test have been quite strikingly confirmed.

The predominantly negative correlations between N scores and conditioning variables are inconsistent with the drive theory of anxiety. The strength of this evidence depends, of course, on the adequacy of N as a measure of anxiety or emotionality. However, in the pilot experiment a negative correlation of -0.26 was found between MAS and rate of conditioning.

The failure to find any significant relationships between E and the conditioning variables is consistent with the results of Willett (1960), who has suggested that it may be the questionnaire measure of extraversion which is at fault rather than the theory relating extraversion and conditioning.

The evidence from other sources for the hypothesis linking introversion-extraversion and excitation-inhibition balance supports this interpretation. However, it should be noted that the present experiment has demonstrated that, under certain circumstances, the correlation between excitatory and inhibitory measures may be positive. Hence some extension of the excitation-inhibition balance hypothesis may be necessary.

Lines of research suggested by the present results include an investigation of non-aversive conditioning in diagnosed schizophrenics. Most studies of conditioning in schizophrenics have used aversive stimuli, and the results have been equivocal.

Peters and Murphree (1954), working with GSR elicited by shock, found that schizophrenics conditioned poorly relative to normals. Mednick (1958), on the other hand, has drawn attention to a number of studies which have all reported superior conditioning of the eye-blink reflex in schizophrenics.

The results of several other experiments are of relevance to the hypothesis that high scores on the Sorting Test are a manifestation of a weak attention response.

Razran (1961) reports a Soviet experiment by Gamburg in which the orienting reactions of normals and schizophrenics were compared.

The sound stimulus used evoked consistent orienting reaction somatic and visceral components in 29 of 34 normals. In the other five normals a defensive reaction was elicited. Only four of 42 schizophrenics gave normal orienting reactions, 15 giving defensive reactions and 23 no reaction at all.

Paintall (1951) compared the GSR of 650 normals and 650 schizophrenics to both shock and verbal threat of shock. He found no difference in response to shock but the responses of the schizophrenics to threat of shock was markedly lower than those of the normals.

Comparatively straight learning curves similar to that of the high scorers on the Sorting Test in non-aversive conditioning have been obtained from schizophrenics by O'Connor & Rawnsley (1959) and Claridge (1960).

O'Connor & Rawnsley studied the performance of schizophrenics on simple industrial sorting and packing tasks over a period of weeks, and Claridge compared the pursuit rotor learning of normals and schizophrenics.

It might be argued that the learning required in these studies differs too much from conditioning to permit meaningful comparisons, but it is conceivable that similar mechanisms may underly the three sets of observations.

REFERENCES

BECKER, W. C. (1960). Cortical Inhibition and extraversion-introversion. *J. abnorm. soc. Psychol.* **61**, 52-66.

BINDRA, D., PATERSON, A. L. & STRZELECKI, J. (1955). On the relation between anxiety and conditioning. *Canad. J. Psychol.* **9**, 1-6.

CLARIDGE, G. (1960). The excitation-inhibition balance in neurotics. In *Experiments in Personality* (ed. H. J. Eysenck). London: Routledge & Kegan Paul.

DAVIS, R. C., SIDDONS, G. F. & STOUT, G. L. (1954). Apparatus for recording autonomic states and changes. *Amer. J. Psychol.* **67**, 343-352.

EYSENCK, H. J. (1957). *The dynamics of anxiety and hysteria*. London: Routledge & Kegan Paul.

FRANKS, C. M. (1957). Personality factors and the rate of conditioning. *Brit. J. Psychol.* **48**, 119-126.

HEATH, H. A. (1961). An empirical study of correlation involving a half-normal distribution. *Psychol. Reports*, **9**, 85-86.

HILGARD, E. R., JONES, L. V. & KAPLAN, S. J. (1951). Conditioned discrimination as related to anxiety. *J. exp. Psychol.* **42**, 94-99.

KAADA, B. R. & JOHANNESSEN, N. B. (1960). Generalized electrocortical activation by cortical stimulation in the cat. *EEG Clin. neurophysiol.* **12**, 567-573.

LOVIBOND, S. H. (1953). Conceptual thinking and schizophrenia. Honours Thesis, Department, of Psychology, University of Melbourne.

LOVIBOND, S. H. (1954). The Object Sorting Test and conceptual thinking in schizophrenia. *Austral. J. Psychol.* **6**, 52-70.

LOVIBOND, S. H. (1961). Aversive and non-aversive conditioning of the GSR. Paper read to the Annual Conference of the British Psychological Society, Australian Branch, Sydney.

MEDNICK, S. A. (1958). A learning theory approach to research in schizophrenia. *Psychol. Bull.* **55**, 316-327.

O'CONNOR, N. & RAWNSLEY, K. (1959). Incentives with paranoid and non-paranoid schizophrenics in a workshop. *Brit. J. med. Psychol.* **32**, 133-143.

PAINTAL, A. S. (1951). A comparison of the galvanic skin responses of normals and psychotics. *J. exp. Psychol.* **41**, 425-428.

PAVLOV, I. P. (1927). *Conditioned Reflexes*. (Trans. G.V. Anrep). Oxford University Press.

PETERS, H. N. & MURPHREE, O. D. (1954). The conditioned reflex activity in the chronic schizophrenic. *J. clin. Psychol.* **10**, 126-130.

RAZRAN, G. (1961). The observable unconscious and the inferable conscious in current Soviet psychophysiology. *Psychol. Rev.* **68**, 81-147.

SPENCE, K. W. (1953). Learning and performance in eyelid conditioning as a function of intensity of the UCS. *J. exp. Psychol.* **45**, 57-63.

SPENCE, K. W. & FARBER, I.E. (1953). Conditioning and extinction as a function of anxiety. *J. exp. Psychol.* **45**, 116-119.

SPENCE, K. W. & TAYLOR, J. A. (1951). Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.* **42**, 183-188.

STEWART, M. A., STERN, J. A., WINOKUR, G. & FREDMAN, S. (1961).* An analysis of GSR conditioning. *Psychol. Rev.* **68**, 60-67.

WILLETT, R. A. (1960). Measures of learning and conditioning. In *Experiments in Personality* (ed. H. J. Eysenck). London: Routledge & Kegan Paul.

* Stewart, Stern, Winokur & Fredman (1961) believe that only GSR's with latencies outside the latency range of unconditioned responses to the CS can legitimately be termed conditioned responses. It has been argued elsewhere (Lovibond, 1961) that this criterion is of relevance only when a DS is not used; if a DS is used it affords a built in control against which the criterion of condition can be validated. In the present experiment the difference between total response to the CS and total response to the DS was significant for all groups.

Manuscript received 30 December 1961

Base-Rate Expectancies and Perceptual Alterations in Hypnosis

By THEODORE R. SARBIN AND MILTON L. ANDERSEN*

University of California, Berkeley

Criticism is made of the hypothesis that hypnotic induction or the 'hypnotic state' is *necessary* to elicit many of the behaviours usually regarded as 'hypnotic phenomena'. Support for this criticism is taken from recent research and reviews in the hypnosis literature. An experiment is reported which was a replication of an experiment in which hypnosis was utilized, where the production of visual hallucinatory behaviour was attributed to the effects of the 'hypnotic state'. The replication experiment did *not* involve hypnosis, but demonstrated hallucinatory behaviour similar to that obtained in the original experiment. The similarity of the results of the two experiments is discussed in terms of: (1) Subject selection and base-rates, and (2) Differences in instructions and motivation given to control subjects and to hypnosis subjects.

INTRODUCTION

A major issue in the area of hypnosis research concerns the necessity of the hypnotic induction procedure (and, by implication, the 'hypnotic state' or 'trance') to bring about many of the behaviours and perceptions usually regarded as 'hypnotic phenomena'. Criticism of the 'necessity hypothesis' in theories of hypnosis has come from three sources: (1) recent research, (2) recent reviews in the hypnosis literature, and (3) role theory.

Some of the recent experiments in hypnosis have directly tested the necessity hypothesis. In a study by Barber (1961a), non-hypnotized persons (people who happened to walk by his office one morning) were utilized as experimental subjects. They were asked to 'concentrate away from red and green' when responding to the plates of the Ishihara Test of Color-Blindness. The responses of these subjects were compared to the responses of a group of deeply hypnotized subjects in a study by Harriman (1942) whose subjects had been given elaborate suggestions to induce colour-blindness. The results were that Barber's non-hypnotized subjects gave as many 'colour-blind' responses on the Ishihara Test as did Harriman's deeply hypnotized subjects. In an excellent study by Fisher (1954) it was clearly shown that post-hypnotic behaviours were elicited only when the subjects believed that the hypnotist-subject relationship still existed. In addition, if the relationship had been broken, the reinstatement of the relationship would again permit the elicitation of post-hypnotic responses. In a study by Orne (1959), one part of the experiment tested the limits of physical endurance (holding a kilogram weight at arm's length) under both waking and hypnotic conditions. However, unlike many previous experimenters, Orne took great pains to insure that adequate degrees of motivation to perform were induced under both conditions. His subjects were able in the waking condition to equal or even exceed their performances in the

* This author was supported in the research by Predoctoral Research Fellowship (MH-13,076-13), from the National Institute of Mental Health, U.S. Public Health Service.

hypnotic condition. In a study by Sutcliffe (1961) the differences in behaviour between the hypnotized and the 'acting' subjects were minimal. He concluded: 'The difference between hypnosis and acting might best be sought elsewhere than in overt behavior . . .' (p. 200).

The second source of criticism of the necessity hypothesis comes from three reviews in the literature about the effects of hypnotic stimulation. In a review by Barber (1961b) on the physiological effects of hypnosis, it is shown that many of the phenomena usually attributed to the effects of the 'hypnotic state' can be produced without any hypnotic induction. In a review by Sutcliffe (1960), entitled: "Credulous" and "Sceptical" Views of Hypnotic Phenomena', a skeptical view is expressed towards attributing any behaviour elicited in a hypnosis situation to the effects of the 'hypnotic state', without also testing the effect in an appropriate experimental design. Sutcliffe proposes such a design, which includes several types of control subjects. In addition, he emphasizes that very few investigators in hypnosis research have employed proper controls and research designs: ' . . . More often than not studies which are held to provide evidence for the "credulous" point of view prove to be inconclusive. The occasional well designed study tends to refute the "credulous" view; and so all told the evidence can be said to directly support or be consistent with the "sceptical" point of view' (p. 97).

The review by Sarbin (1955) on the 'Physiological Effects of Hypnotic Stimulation' presented an application of the concept of 'organismic involvement' to hypnotically-induced alterations of physiological processes. In role-theoretical terms, organismic involvement is one dimension of role-enactment. The behaviour of the person enacting the role of the hypnosis-subject can be regarded as occurring at some level of involvement. At high levels of involvement, the intense striving of the hypnotic subject is observed in the rapid pulse, increased respiration-rate, grimacing, enhanced muscular contractions, etc. Intensified activity in the skeletal musculature brings about a concomitant increase of activity in the organs served by the autonomic nervous system. With this concept of organismic involvement as a mediating process, several studies were reviewed in which physiological changes were preceded by hypnotic stimulation. It was demonstrated that, in the several experiments reviewed, the production of physiological changes may be more appropriately attributed to the increased involvement of the hypnosis-subject than to the 'trance-state' of hypnosis. An obvious implication of the review is that properly motivated non-hypnotized subjects would be able to parallel the performance of hypnotized subjects.

The third source of criticism of the necessity hypothesis stems from role-theoretical concepts: Sarbin (1943, 1950, 1954, 1955). Role theory emphasizes the obvious social psychological nature of the hypnosis situation, and calls for concepts which are continuous with social psychological formulations. The behaviour of the hypnosis-subject is regarded as one form of a more general form of social psychological behaviour—role enactment. The effectiveness of role enactment is taken to be a function of: (1) favourable motivation (congruence of self-concept with the role of hypnosis-subject), (2) appropriate role perceptions, and (3) role skills or aptitude. The role enactment itself can be meaningfully described

in at least two dimensions: (1) Accessibility to self-report, and (2) Degree of organismic involvement. The phrase, 'continuous with social psychological formulations', is to emphasize that so-called 'hypnotic phenomena' are not unique to the hypnosis situation. Furthermore, the variables of which hypnotic behaviour is a function (motivation, perception and aptitude), are also determinants of other behaviours. These other behaviours would be those elicited in various social-influence situations: persuasion, conformity, attitude change, suggestion, and hypnotic-like situations. These role-theoretical concepts are intended to sensitize the hypnosis researcher to the ubiquity of general social psychological processes, and to restrain him from neglecting them in his experimental designs and theoretical attempts.

The experiment reported in this paper is a partial replication of one of the many experiments in the hypnosis literature which attribute the results obtained to the effects of the 'hypnotic state'. This experiment, Underwood (1960), was chosen to be replicated because its careful execution permitted identification of its ambiguities and weaknesses (which Underwood himself indicated). Underwood's experiment utilized hypnotic induction, while the replication experiment did not. The intent of the replication experiment was to determine if the performance of Underwood's hypnotized subjects could be paralleled by non-hypnotized subjects.

UNDERWOOD'S EXPERIMENT

Subject selection

Underwood utilized three groups of subjects in his experiment. Subjects in two of the groups were hypnotized during the experiment; subjects in the control group were not hypnotized. Because of the extreme importance of subject selection to the argument of this paper, we quote from Underwood *in extenso*:

Hypnotic Ss were selected by screening approximately 195 Ss (mostly graduate students) for hypnotic susceptibility by attempting to hypnotize them . . . Ss who showed susceptibility in either a group or individual session were seen privately for several sessions to determine whether or not they would be capable of achieving sufficient trance depth to hallucinate. Twelve Ss were chosen as being suitable for the study on the basis of seeming to have the most vivid hallucinations, while lacking in any gross visual defects or any background that might have made them familiar with the illusions. They all responded positively to trance depth tests of eyelid and arm catalepsies, arm levitation, hand anaesthesia, selective amnesia, positive and negative tactial, auditory and visual hallucinations, and post-hypnotic suggestions, including trance induction by signal. Also, they were all somnambulists; that is, they could walk, talk, and have their eyes open while remaining in the trance.

Even within this highly select group, six Ss showed imperfections in their trance performance. Three reported hallucinations that were sometimes 'ghost-like' or fluctuating in vividness, and they also reported that hallucinations required considerable effort. Three other Ss reported very vivid hallucinations, but were obviously not hallucinating with ease. There was much squinting, tear secretion, fatigue and other signs of excessive effort. The remaining six Ss reported constantly vivid hallucinations (indistinguishable from reality) which were apparently produced with ease.

... Those six Ss who showed trance imperfections were delegated to a separate group called the 'deep group'. The six who showed flawless performances were called the 'very deep group'.

The control group of six Ss was picked to match the hypnotic groups roughly in age and educational background. No member was to have ever been hypnotized ... (Underwood, 1960, pp. 39-40).

This lengthy passage from Underwood is to emphasize that the major criterion for the inclusion of a subject in his 'very deep group' is his ability to produce vivid hallucinations. The perceptual performances of these hypnotic Ss were then to be compared with non-hypnotized control subjects on an experimental task that called for vivid hallucinatory ability.

Procedure

He utilized as a task of 'positive hallucination' a figure-ground (figure-field) illusion presented on a translucent glass screen, in which the presence of the field distorts the

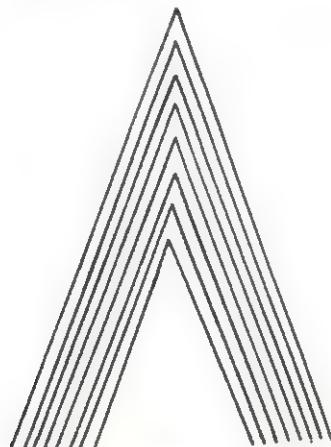


Fig. 1.

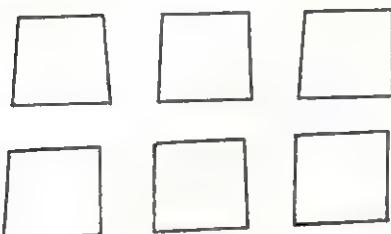


Fig. 2.

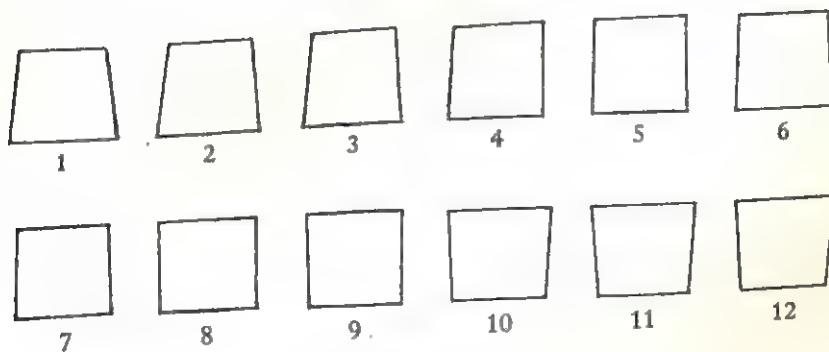


Fig. 3.

appearance of the figure. He reasoned that effective hallucination of the field would distort the appearance of the figure in the same manner as does the *actual presence* of the field. As stimulus material he used a pattern of inverted V's as the field (Fig. 1), and each one of a series of six squares and near-squares (trapezoids) were used as the figure (Fig. 2). The subject's

task upon each presentation of one of the figures was to indicate how the figure appeared to him. This was accomplished by having within view at all times a set of 12 comparison figures on a large card (Fig. 3). The subject then indicated which one of the 12 comparison

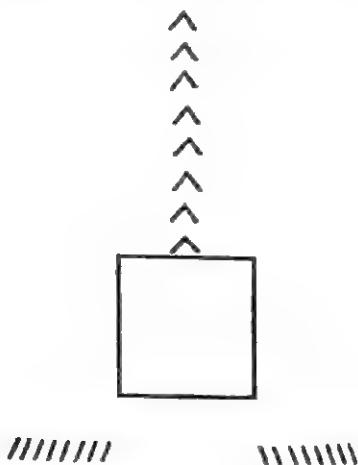


Fig. 4.

figures appeared to him as the best match of the figure being presented. (The six figures that were individually presented on the glass screen have the same dimensions as comparison figures 3, 4, 5, 6, 7 and 8.) Underwood chose to employ this particular illusion because it is rarely known, even among psychology majors, and would therefore be difficult to fake one's responses or to try to please the experimenter.

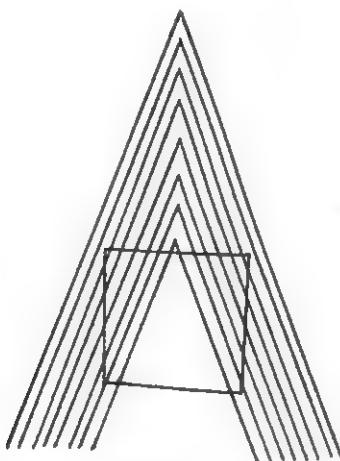


Fig. 5.

The stimulus material was presented under three conditions in the following order:^{*} (1) Under 'objective condition' each figure was presented alone, that is, *without* the field, to determine the 'base-line' appearance of the figures. An example is Fig. 4. The little marks are 'reference marks' to indicate where the inverted V's composing the field would be located later under the 'illusion condition'. (2) Under 'hallucination condition' each figure

* Only a portion of Underwood's experiment is under discussion in this paper. It is that part which gave him positive results, and which was employed in our replication study. Certain terminology used here deviates from that used by Underwood. For example, the 'objective condition' described here are actually his 'reference condition'. These changes have been made for clarity of discussion, and do not alter the argument of this paper.

is again presented without the field. Figure 4 therefore again serves as an example of the presentation of one figure. However, there are now instructions to the hypnotized subjects to *hallucinate* the presence of the field where the little marks indicate it should be, and with instructions to the control Ss to *imagine* the presence of the field. (3) Under 'illusion condition' a figure is presented with the field superimposed upon it. Figure 5 is an example. The effect of the presence of the field (the inverted V's) is to make the figure appear wider at the top.

Results

Underwood presented his results graphically in terms of 'mean error per figure-judgment' with the error in figure-units. For example, if a subject upon presentation of a figure on the glass screen reports that it appears to him as No. 4 comparison figure, but it actually has the dimensions of comparison figure No. 3, his

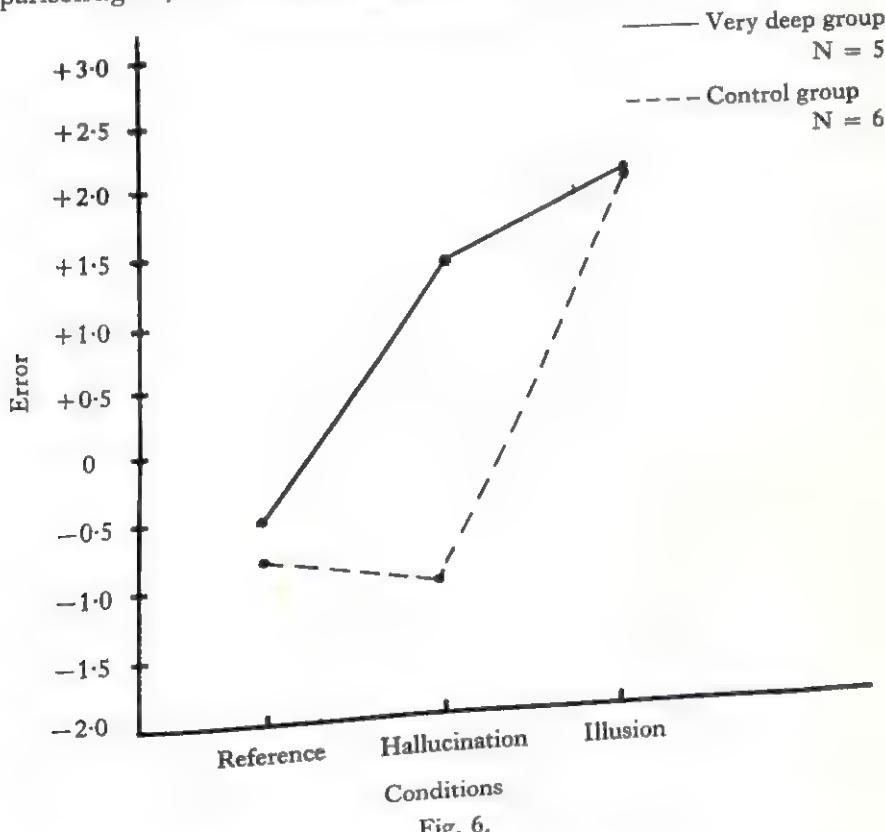


Fig. 6.

error for that figure judgment is +1. Or if, upon presentation of figure No. 3 a subject reports that it appears to him as figure No. 2, his error for that figure judgment is -1.

In reproducing the graph used by Underwood, two groups are omitted. These are: (1) His 'deep group'. The performance of this group did not appreciably or significantly differ from the performance of his control group. (2) His *total 'very deep group'*, numbering 6 subjects. One subject in this group showed erratic responses, his responses deviated from those of all other Ss, and had once been hospitalized for a hysterical reaction. He was therefore excluded from the

'very deep group', and the scores of this group recalculated from the responses of the remaining 5 members. The scores of his 'very deep group' presented here are *without* the scores of this one deviant subject. The amended graph is presented as Fig. 6. The difference between the 'very deep group' and the control group under the hallucination conditions is significant at the 0.005 level, by a one-alternative *t* test.

BERKELEY EXPERIMENT

Selection of subjects

The subjects used were 120 students at the University of California in Berkeley. There was no selection of subjects, other than that of availability to the experimenter. Five groups of subjects were used: Two sections of an introductory course in psychology, two advanced undergraduate classes in psychology, and a University extension class in psychology for civilian management personnel of a U.S. Navy base. There were no subjects or groups utilized as 'pilot studies', 'dry-runs', or 'try-outs'. The results to be presented are taken from *all* the subjects tested in the experiment.

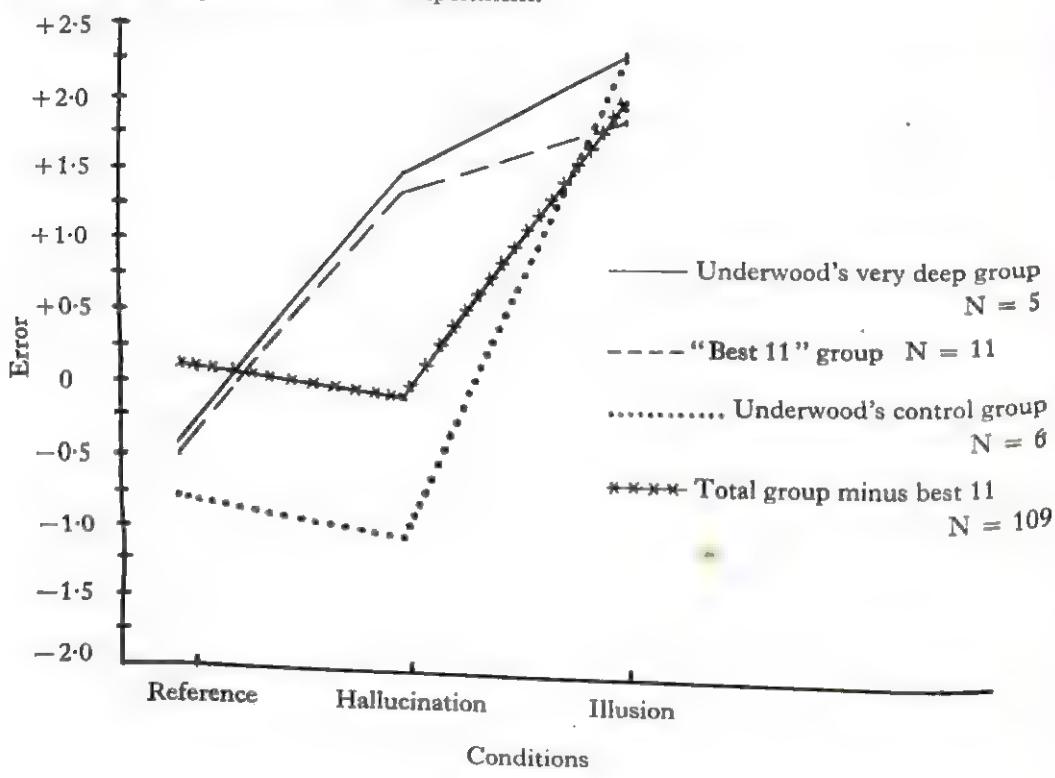


Fig. 7.

Procedure

The stimulus materials and order of presentation were the same as in that portion of the Underwood study reported above. However, there were four procedural differences between the two experiments. (1) The data were gathered in group sessions instead of individually. (2) The figures were not projected on to a translucent glass screen, but were presented on large white cards (22 in. \times 28 in.) to three groups of *Ss*, and by slide-projector to two groups of *Ss*. (3) In the Berkeley experiment there was no utilization of 'hypnotic induction' or a 'trance'. All subjects were run in the waking state. (4) In all the groups tested in this experiment, the *Ss* were given careful and detailed instructions designed to enlist their co-operation, to arouse their interest, and to enhance and support their imagination during the 'hallucination condition'. That is, the subjects were not merely asked to guess what the

non-present field might look like, but to strive intensely to bring about a vivid, imaginative, and life-like facsimile of the non-present field.

Results

The results of this experiment are presented in Fig. 7, together with the results of Underwood's very deep group (minus the deviant *S*), and his control group. The two groups from the Berkeley study were created in the following manner: Out of the total *N* of 120 subjects, the 11 subjects whose scores indicated the most effective hallucinatory (or imaginative) behaviour during the hallucination condition, formed the 'best 11' group. The remaining 109 subjects formed the 'total group minus best 11'.

DISCUSSION

There are three main points that require discussion.

(1) *Evaluation of the similarity of the results from the two experiments.* From a visual inspection of the graph (Fig. 7), it is apparent that the data are highly similar, but the problem of evaluating this similarity remains. In the replication study the '11 best' group was simply those 11 subjects scoring at the top of the distribution of scores. It did not seem proper or meaningful, therefore, to compute a statistic in order to compare the means of these two groups, or to make a formal statistical comparison of the results of the two experiments. It seems sufficient to demonstrate that in a population of university students, 11 out of 120 *Ss* (or about 9 per cent in the Berkeley study) could produce the behaviour obtained by a 3 per cent sample (six *Ss* out of 195 in the Underwood study).

Another problem in the data that must be handled is: Could the results of the Berkeley study be due merely to chance? This is difficult to evaluate, for no additional measures were obtained to provide an estimate of reliability. A measure of reliability is difficult with this particular illusion task, for the acquaintance gained with the illusion in one study destroys its usefulness in a second study using the same *Ss*. 'Comparable' illusion tasks could be utilized. However, the portions of Underwood's experiment not discussed in this paper concern the use of three similar illusion tasks, all with negative results. If these other three illusion tasks are to be regarded as 'comparable forms' of the illusion task with which Underwood did obtain positive results, he did not, then, demonstrate a reliable phenomenon.

(2) *Subject selection and base-rates.* Underwood's own description indicates that a major criterion in his extensive and rigorous selection of subjects for his hypnotic groups was their ability to produce vivid hallucinations. It seems reasonable, therefore, that any differences obtained between a carefully selected hypnotic group and a control group could be attributed to selecting out the 'good hallucinators' from the general population of subjects. As already stated, the 'best 11' group (9 per cent of one population) was similar in its hallucinatory behaviour to the 'very deep' group (3 per cent of another population).

In the usual procedure for testing responsiveness to hypnotic induction, multiple tests are used, such as, speed of eye closure, catalepsy, post-hypnotic performances,

amnesia and so on. Thus Underwood's subjects had to pass multiple criteria (including ability to hallucinate) in order to belong to the 'very deep group'. Had he used a single criterion, as we did, then his 'very deep group' might well have comprised 9 per cent of the total group he screened for hypnotic susceptibility.

Common to many studies that report changes in perceptual or somatic function, as a result of hypnotic induction procedures, is the error of assuming zero base rates for phenomena that have low but reliable base rates. The need is urgent to examine carefully base rates for any test that are employed to validate the conception of a trance state and its supposed effects. This caution is particularly pointed when we pre-test large numbers of persons on tests that are correlated with tests used in the criterion measures. Elsewhere we have shown how the proportion of subjects passing specific items in hypnotic susceptibility tests is a function of the base rates obtained without hypnosis from unselected subjects (Andersen and Sarbin, 1963, in press).

(3) *Differences in instructions and motivation given to control subjects and to hypnosis subjects.* Underwood's instructions to his control Ss under the 'hallucination condition' were: 'Instructions to control Ss . . . were the same for hypnotic Ss, except that they were told to *imagine* the lines being filled in . . . and they were to *guess* the forms the projected figures would appear to have, if the suggested stimulus condition actually existed.' It is evident that they were not led to expect that they would be able to perceive the non-existent field, as the hypnotic Ss were given every inducement to so expect. Furthermore, what he refers to as 'minor differences in wording' would be crucial differences when the wording maintains the expectation of the control Ss that they can only guess, while the wording to the hypnotic Ss maintains their expectations that they will, and in fact, *should* perceive the non-existent field.

REFERENCES

ANDERSEN, M. L. & SARBIN, T. R. (1963). Base rates expectancies and motoric alterations in hypnosis. In press.

BARBER, T. X. & DEELEY, D. C. (1961a). Experimental evidence for a theory of hypnotic behavior: 1. 'Hypnotic color-blindness' without 'hypnosis'. *Int. J. clin. exp. Hypnosis*, **9** (2), 70-86.

BARBER, T. X. (1961b). Physiological effects of 'hypnosis'. *Psychol. Bull.* **58** (5), 390-419.

FISHER, S. (1954). The role of expectancy in the performance of post-hypnotic behavior. *J. abnorm. (soc.) Psychol.* **49**, 503-507.

HARRIMAN, P. L. (1942). Hypnotic induction of color vision anomalies: I. The use of the Ishihara and the Jensen tests to verify the acceptance of suggested color blindness. *J. genet. Psychol.* **27**, 289-298.

ORNE, M. T. (1959). The nature of hypnosis: artifact & essence. *J. abnorm. (soc.) Psychol.* **58** (3), 277-299.

SARBIN, T. R. (1943). The concept of role-taking. *Sociometry*, **6**, 273-285.

SARBIN, T. R. (1950). Contributions to role-taking theory. I. Hypnotic behavior. *Psychol. Rev.* **57**, 255-270.

SARBIN, T. R. (1954). Role-theory. In *Handbook of social psychology* (ed. G. Lindzey). Cambridge, Mass.: Addison-Wesley Pub. Co.

SARBIN, T. R. (1956). Physiological effects of hypnotic stimulation. In *Hypnosis and its therapeutic implications* (ed. R. M. Dorcus). New York: McGraw-Hill.

SUTCLIFFE, J. P. (1960). 'Credulous' and 'skeptical' view of hypnotic phenomena. *Int. J. clin. exp. Hypnosis*, 8 (2), 73-102.

SUTCLIFFE, J. P. (1961). 'Credulous' and 'skeptical' view of hypnotic phenomena. *J. abnorm. (soc.) Psychol.* 62 (2), 189-200.

UNDERWOOD, H. W. (1960). The validity of hypnotically induced visual hallucinations. *J. abnorm. (soc.) Psychol.* 61 (1), 39-46.

Manuscript received 21 October 1962

Individual Vocabulary and Semantic Currency: A Preliminary Study

BY R. C. OLDFIELD

Institute of Experimental Psychology, Oxford

A vocabulary test, whose items are chosen by random sampling of a large dictionary and which can therefore afford estimates of vocabulary on an *absolute* basis, is described and the data obtained on a group of 40 subjects are reported. The test is shown to have a high reliability as assessed by a split-half criterion. Its application to the study of 'semantic currency'—the diffusion of knowledge of word-meanings—in a social group is illustrated, and some implications for the analysis of some aspects of the cultural structure in a community are discussed. The possibility of deriving a short test capable of rapid and wide application is considered.

INTRODUCTION

Any highly evolved language, such as modern English, possesses a very large number of words, and no single person knows the meaning of all of them. There are many reasons for this—the accretion of archaic forms, the growth of special, esoteric and technical vocabularies, the progressive refinement of semantic differentiation, uneven distribution of education and intelligence, the segregation of social classes and so forth. But the distribution of vocabulary among individuals and the effective currency of words—that is, the extent to which knowledge of their meanings is widespread or limited in a community—deserve closer and more quantitative study than they have hitherto attracted.

In the first place our contemporary society is often said to be beset by failure in communication and sharing of ideas between different sections of the community. But views about this are largely innocent of any objective criteria or empirical determination of its character and extent, other than those afforded by opinion and attitude surveys. These are subject to a number of uncertainties, in particular those resulting from the fact that such techniques themselves depend on assumptions about the efficacy of verbal communication. Mr M. Argyle has provided me with an excellent example of this from his own experience. Using the Maudsley Personality Inventory (which the subject fills in himself) in connection with a study of Borstal inmates, he discovered that these boys, aged 16 and upwards, quite commonly did not know the meaning of a number of words used in the questions of the test. Among such words were 'initiative', 'moody', and 'touchy'. The word whose meaning was most often not known was 'disgruntled'.

Nobody would suggest that the question whether an individual possesses and can use a given concept or idea can be completely settled by finding out whether he knows the meaning of the word corresponding to it. But it must be acknowledged that unless he *does*, communication in terms of the concept or idea would be difficult, if not impossible, and—at worst—misleading. Our society is increasingly the scene of exercises in 'mass-communication', a considerable part of which are

verbal. Yet little attempt has been made to find by any simple, objective means how far such communication is, in fact, limited or ineffective by restricted knowledge of word-meanings among those to whom it is addressed. That verbal material is accepted—and often paid for—in large quantities is no evidence as to the consumer's ability to 'benefit' by it. If, therefore, it can be shown that reliable estimates of individual vocabulary and of semantic currency can be made on a statistical basis, this might provide a useful weapon for an attack on a number of important problems in the study of society.

Another, more immediately practical, ground for undertaking such determinations lies in the recent immense development in the learning of certain languages, particularly English, as a means of communication among people of other mother tongues. This has raised problems in the techniques and economics of teaching which need scientific attack. A quantitative estimate not only of the existing linguistic resources of English, but also of the extent to which in different circumstances these are, or can be, drawn upon in practice would seem to be a matter of some urgency.

A third reason for undertaking the work reported in this paper—and it was this which originally suggested the present investigation—is connected with recent developments in the field of experimental linguistics, especially those stemming from the application of information theory. Here, techniques involving word-guessing or word-choice are often used. So, for instance, the methods for the construction of statistical approximations to English, originated by Shannon (1948), developed by Miller & Selfridge (1950) and further employed by Moray & Taylor (1959) rest upon the selection by a number of individuals of a word appropriate to a given context. To some degree, therefore, the results must be dependent on the verbal repertoire of the subjects employed. Indeed, the last-named authors have noted (private communication) certain differences between results obtained by this technique in this country and in the United States. Again, Shannon (1951) used a guessing technique in his estimation of the redundancy of English prose. Treisman (1962) following Taylor (1955), developed a simplified form of this artifice and 'calibrated' it by applying it to passages of known statistical order of approximation. Here, again, results might be expected to depend upon the vocabulary resources of the guessers. It is, of course, clear in these connections that the *organization* of the vocabulary store is at least as important as its *total* contents. Nevertheless, as a first step in trying to analyse this organization, it seemed appropriate to make some kind of estimate of the total number of items the store contains.

Doubts may be felt about whether in practice it is possible to find out by simple means whether or not a given individual knows the meaning of a given word. But for many years now the testing of an individual's knowledge of the meaning of words has been a serviceable device in psychometric practice. Such determinations are, of course, on a relative, not an absolute basis, since they are made by the use of a selected and graded series of words. The principle of the present investigation, on the other hand, is to use a series of words randomly selected from the total vocabulary of the language. This principle was used by Smith (1941), following Seashore & Eckerson (1940), to determine the absolute vocabulary of school-

children and its variation with age, but for a number of reasons this work is not closely comparable with the present investigation.

Suppose we have a list of N_T words randomly chosen from the total number N_L of words in the language. (Some limitations imposed by substituting, for N_L , N_D —the number of words in a large dictionary—are noted below.) With this list we test S individuals who belong to a total population. The S individuals may be chosen so as to belong to some class or subgroup definable in respect of some variable such as age, socio-economic status, or educational category, or they may form a representative sample of the whole population. Empirically we find whether an individual s_j knows the meaning of a word w_k and thus obtain a matrix

$$M = (m_{jk}) \quad (1)$$

in which the value of m_{jk} is 1 or 0 according to whether s_j does or does not know the meaning of w_k . Then

$$V_j = \frac{1}{N_T} \sum_{k=1}^N m_{jk} \quad (2)$$

is individual s_j 's '*test vocabulary*'—that is, the number of test words whose meaning he knows expressed as a fraction of the number of words used in the test. Subject to the usual conditions of random sampling this is also the proportion of the total number of words whose meaning he knows and his '*absolute vocabulary*' is therefore

$$V'_j = N_L \cdot V_j \quad (3)$$

or writing N_D for N_L

$$V'_j = N_D \cdot V_j \quad (4)$$

Similarly

$$C_k = \frac{1}{S} \sum_{j=1}^S m_{jk} \quad (5)$$

is the word w_k 's '*semantic currency*'.

METHOD

1. *Estimation of N_L —the total number of words in the language.* We take N_D —the number of words in a large dictionary—to be an estimate of N_L . There are obvious objections to this. Some words, for instance, will not be entered in any dictionary, however large. These words, however, are likely to be ones whose meaning is unknown to any but a very small proportion of the population. The larger the dictionary, the larger will be the proportion of obscure words and hence, for a randomly selected test series of *constant* length, any individual will have a *smaller* test vocabulary the larger the dictionary. Whether or not this will correspond with a larger or smaller absolute vocabulary will depend on whether the value of

$$\left[\frac{\partial V_j}{\partial N_D} \right]_{N_T}$$

turns out to be greater or less than unity. The data of Seashore & Eckerson (1940) suggest, as does Hartmann (1941), that the larger the dictionary used the larger the individual absolute vocabulary estimate will be. However, since we are interested in individual differences of absolute vocabulary and in their distribution in a population, the point of practical importance is that the dictionary chosen should not, by reason of its smallness,

materially limit the variance of the test results. The *Oxford Shorter English Dictionary* (1933) was used and seemed to meet the needs of the investigation.

This contains 2475 pages, out of which 100 were selected by use of Kendall & Babington Smith's (1938) random sampling numbers. The number of words—or more properly *word-entries*—on each of these pages was counted. The criteria according to which anything was accepted or rejected as an independent word-entry were kept simple, since these same criteria were applied in selecting words for inclusion in the test list. Thus, for example, foreign words—chiefly Latin—were accepted if they were considered by the author to have general currency in the English language. If, on the other hand, they were purely technical—e.g. legal—they were rejected. Compound entries, if corresponding to some special entity, such as *blackleg*, were accepted, while those having scarcely more semantic implication than would be suggested by their component parts—such as *blackcurrant*—were rejected. In practice, the application of criteria such as these did not present any great difficulty, as has been noted by previous workers.

2. *Selection of words for test.* Two hundred words were picked out of the dictionary, Kendall and Babington Smith numbers being used to determine both the page and the number of the entry on each page to be chosen. The same criteria as to word-entries were employed. The words thus selected were typed in capitals on plain 5 in. \times 3 in. index cards and these were arranged in alphabetical order. Since the actual test series is likely to be used again, it is not disclosed in this paper. But in Appendix A is given a similar randomly chosen series of 50 words.

3. *Administration of test.* The whole series was presented one by one to each of 40 subjects and he or she was asked to state what the meaning of the word in question was. The criteria used to decide whether or not the meaning was known were of the kind used in the classical vocabulary test as included in Wechsler-Bellevue procedure. The subject was given every chance to make himself clear. Where a word had more than one meaning, any one correctly defined was accepted, as were acknowledged guesses if correct. The subjects being highly educated, lengthy discussion sometimes developed, and where some doubt still remained a half point was conceded by way of bringing this to an end. These half points were subsequently deleted.

4. *Subjects.* Since the original aim of the investigation was to get a figure for the total semantic resources possessed by individuals of the kind who had been subjects in word-guessing and word-choice experiments—especially those carried out in this Institute—the same population was drawn upon. Of the 40 subjects used, 30 were male and 10 female; their ages ranged from 19 to 60, with a mean of 29.5 years. Sixteen were undergraduates, 17 graduates and 7 neither. All the latter had been educated to G.C.E. ordinary level or beyond.

RESULTS

1. Number of words in dictionary

Number of pages in dictionary = 2475.

Mean words per page = 44.51 (100 sample pages).

Standard error of mean = 1.94.

Hence total number of words in dictionary = $110,162 \pm 4800$.

Each word in random test list of 200 words therefore represents 551 ± 24 words.

2. Vocabulary in the group tested

Mean score = 135.25. Range 99-162. $\sigma = 11.17$.

Mean test vocabulary = 0.676.

Mean absolute vocabulary = 74.523.

Mean scores of subgroups

(U) Undergraduates ($n = 16$) = 135.56

(G) Graduates ($n = 17$) = 139.77

(O) Others ($n = 7$) = 123.57

Differences of means

$$\begin{array}{ll} U \times G = 4.21 & p > 0.05 \\ U \times O = 11.99 & p < 0.02 \\ G \times O = 16.20 & p < 0.01 \end{array}$$

3. *Split-half test reliability*

Correlation of two sub-tests (alternate items - $n = 100$) $r = +0.855$.

Index of reliability of each subtest $= \sqrt{r} = 0.925$.

From Spearman-Brown formula, reliability of full test $= 0.961$.

Standard error of estimate of true score based on actual score S is (Guilford, 1936, p. 415) $s.e. = \sigma_s \sqrt{0.961(1-0.961)} = 0.195\sigma_s$,

But $\sigma_s = 11.17$, hence $s.e. = 2.18$

$$\text{or } s.e.(V_j) = \frac{2.18}{200} = 0.011.$$

4. *Estimation of standard error of V'_j (absolute vocabulary of subject j)*

From equation (4)

$$V'_j = N_D \cdot V_j.$$

We have, for variance of a product,

$$\text{var. } V'_j = N_D^2 \cdot \text{var. } V'_j + V_j^2 \cdot \text{var. } N_D.$$

$$\text{Now } N_D = 110,162, \text{var. } N_D = 4800^2, \text{var. } V_j = 0.011^2.$$

Substituting these values, Table 1 has been computed giving standard errors for values of absolute vocabulary within the range found in our subjects. It will be seen that the size of the standard error remains substantially constant, in relation to the vocabulary values, at about 5 per cent.

Table 1

Absolute vocabulary estimates	Standard error of estimate	$\frac{s.e.(V)}{V}$
50,000	2494	0.050
60,000	2882	0.048
70,000	3280	0.047
80,000	3689	0.046
90,000	4106	0.046
100,000	4526	0.045

5. *The semantic characteristic*

There are various ways in which our results may be graphically displayed as a whole. We might, for instance, plot the frequency distribution for words of a given semantic currency. This would be U-shaped, since a relatively high proportion of the 200 words were either known to all the subjects (85 words) or to none of them (20 words), while the remaining 95 words were distributed between all the intervening currency values. It is, however, more convenient to use a cumulative version of this frequency distribution. To obtain this we arrange the 200 words in order of currency and plot the currency values as ordinate against the rank order. Where there are tied ranks the centre point is plotted. Our data are represented in

this way in Figure 1, in which the best curve has been drawn by hand through the points. The ordinate scale of semantic currency is expressed both in number of subjects knowing the meaning of each word and also as semantic currency C_k . The abscissa scale is shown not only as rank order 0 to 200 but as proportions of the number of words in the total vocabulary of the language. In interpreting this graph it is important to note that these 200 words were drawn at random from the

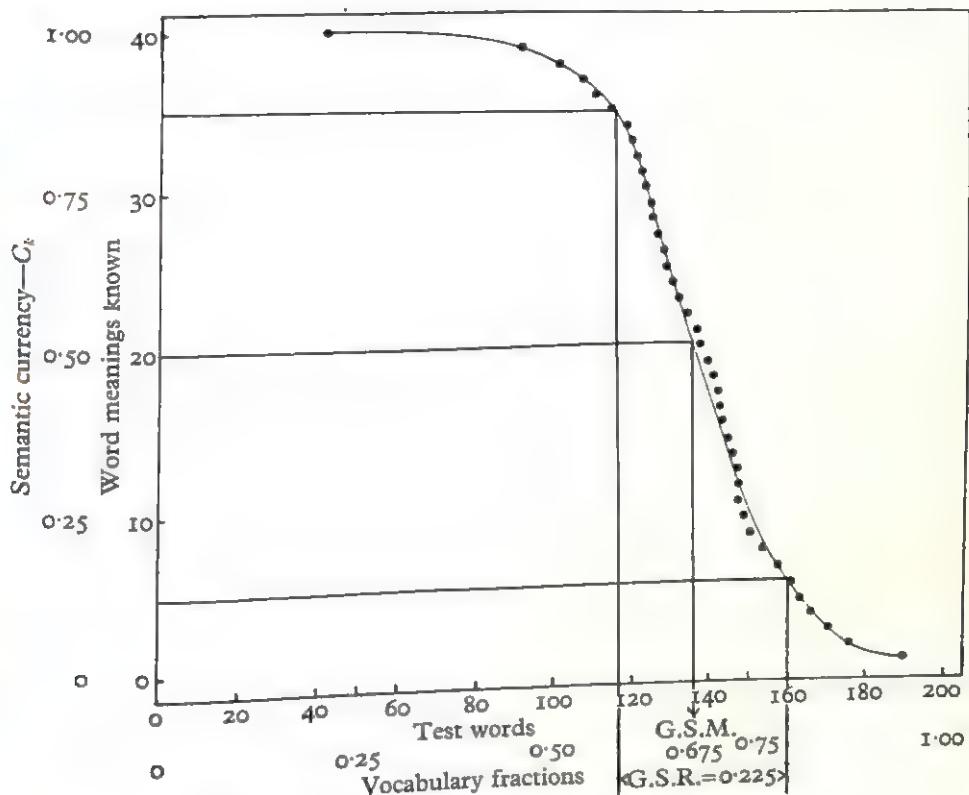


Fig. 1. Semantic characteristic for group investigated.

total words in the dictionary. Hence the abscissa scale is not purely ordinal in character but represents fractions of the total vocabulary of the language.

Such a characteristic may be suitably defined by two parameters: (1) The abscissa value corresponding to a semantic currency of 0.5, and (2) the abscissa range between currency values of (say) 0.125 and 0.875. We may call the first 'the group semantic median' and the second 'the group semantic range'. In this particular case we have G.S.M. = 0.675 and G.S.R. = 0.225.

DISCUSSION

The chief points of interest which emerge from this preliminary study are the following:

(1) It is possible, using a vocabulary test of which the items are chosen by random sampling but whose administration follows the classical lines of the

intelligence test, to establish estimates on an absolute basis of the vocabulary of an individual, and of the extent to which knowledge of the meanings of words is diffused within a given community.

(2) The reliability of the test as indicated by the split-half method is high, the

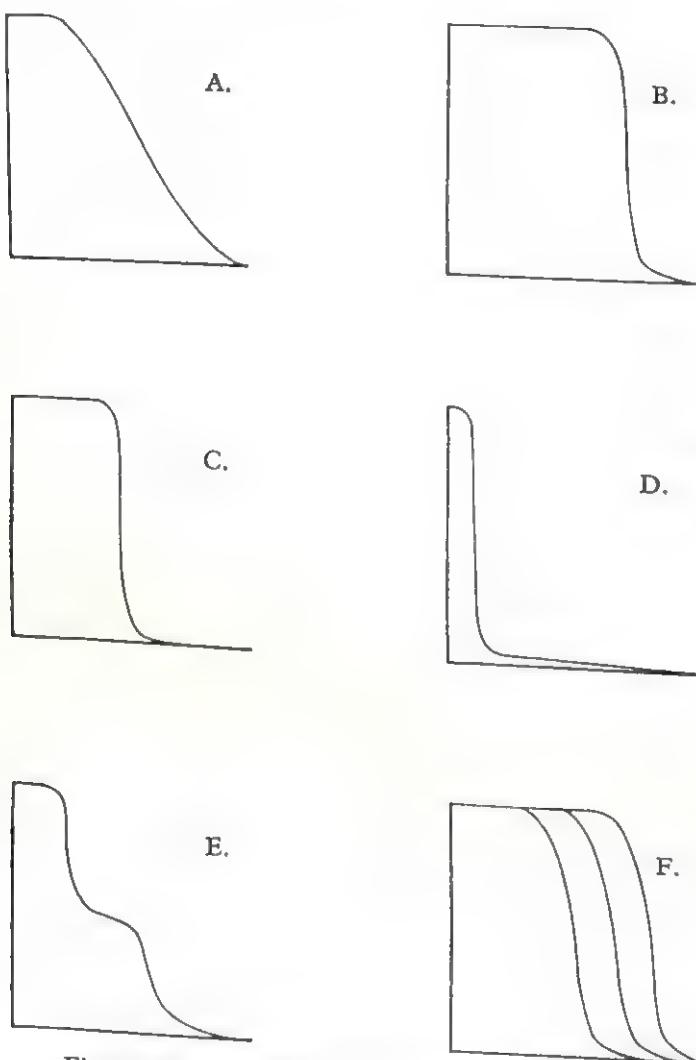


Fig. 2. Hypothetical forms of semantic characteristic.

standard error of the final estimate of an individual's vocabulary being about $1/20$ of the estimate itself.

(3) Even with the small numbers of subjects used, the test proved sufficiently sensitive to show up significant differences between non-university, but educated people on the one hand, and graduates or undergraduates on the other.

The data available to us do not, unfortunately, allow any conclusions to be drawn about the causes of variation of individual vocabulary within a group. One would

certainly expect part of the group variance to be accountable for in terms of variance in I.Q., but measures of this were not available. It would seem possible, too, that, at any rate among individuals of a restricted range of high intelligence, there would be an age effect. One might expect that the meanings of the more recondite words would be acquired with time spent on reading etc. The product-moment correlation for our group, however, falls short of significance at $+0.274$ (for $r = 0.304$, $p = 0.05$).

It is perhaps a permissible speculation to consider what different forms the semantic characteristic will take in different linguistically organized communities. These forms will, of course, depend upon two factors: the sample of the population tested, and the type of semantic diffusion. In the first place, if we obtained data from a large representative sample of the whole community and there were wide variations of vocabulary resources within it, we should expect a characteristic something like that shown in Figure 2A. If, on the other hand, the whole community possessed a very high degree of linguistic uniformity, the characteristic would tend towards that shown in B, if the vocabulary of the language does not incorporate any high proportion of archaic items. If, on the other hand, it does, the characteristic will take the form of C. A community which consisted of a small, exclusive ruling class possessing a high degree of cultural sophistication and linguistic refinement on the one hand, and, on the other, of a large mass of uneducated helots restricted in their mental *ambiance*, would have a characteristic such as that shown in D. Should, on the other hand, a more complex class system prevail, showing, say, three linguistically discriminable elements, the characteristic might be comparable with that shown in E. Finally, if we plotted the set of characteristics for a number of restricted and culturally homogeneous subgroups in the community, we would expect to find a series of curves as in F.

Lastly, a word may be said about the length of the test and the possibility that, data having been obtained using a large number of test items, a much abbreviated form capable of wide and rapid application might be constructed. If, for the present, we confine ourselves to an application of the test to restricted and homogeneous subgroups, not too far removed from that which was the subject of the present investigation, a number (say 20) of words whose rank orders fall within this region could be extracted, spaced at equal intervals on the ordinate, or currency, scale. Such a test could be administered with considerable speed and simplicity and the data resulting from it would enable one to plot the corresponding straight line portions of the characteristics for the subgroups in question. The group semantic means and the group semantic ranges could then be estimated, if necessary by extrapolation. If, however, we wished to have available a series of abbreviated tests suited to the estimation of these parameters for subgroups of varying cultural levels throughout the entire community, it would be necessary first to apply the full test to an adequate representative sample of the entire population, in order to determine the rank order of all, or at least the great majority of, the 200 words.

* Further work is in progress along these lines. I would be very glad to hear from anybody who might like to co-operate in obtaining data.

APPENDIX A

Series of 50 entries randomly selected from the *Shorter Oxford English Dictionary*

ABCISION	ELECTORAL
FUSIFORM	SCREW
RUMLY	RELAXANT
POMEWATER	HYSTEROPHYTAL
WIDENER	ENERGIATYPE
SYNTHERMAL	PINCERS
TRAILING	TRAPPER
BEGGARLY	INEXTRICABLE
SYNDESM	VESUVIAN
SCLEROBASE	UNDISHONoured
REIMBURSE	PEDATED
WET	RAREFACTIVE
UNTRUSTWORTHY	STREAMLESS
DONATISTICAL	COLLODIONIZE
SPARKLING	RECOVEREE
STUFFING	MINUTE
TOG	PURGATORY
STALAGMITICALLY	SCURFY
MAST	VENISON
TYPHOMANIA	CONNIVER
DEFUNCT	PROVERBIALITY
ONERATE	ALABASTRINE
PUNCTURE	REGENERATIVE
SCAR	LEGULEIAN
TACHYDRITE	PERCHANCE

I wish to put on record my most grateful thanks to the late Mrs Adrienne Dunn, who carried out the actual administration of the test and whose death at an early age was a sad loss to those whom she had helped at the Institute of Experimental Psychology.

I have also to thank Messrs B. Babington Smith and J. F. Scott for statistical help.

REFERENCES

GUILFORD, J. P. (1936). *Psychometric Methods*. New York & London: McGraw-Hill.

HARTMANN, G. W. (1941). A critique of the common method of estimating vocabulary size. *J. Educ. Psychol.* 32, 351-8.

KENDALL, M. G. & BABBINGTON SMITH, B. (1938). Randomness and random sampling numbers. *J. R. statist. Soc.* 101, 147-166.

MILLER, G. A. & SELFRIDGE, J. A. (1950). Verbal context and the recall of meaningful material. *Amer. J. Psychol.*, 63, 167-185.

MORAY, N. & TAYLOR, A. M. (1960). Statistical approximations to English and French. *Lang. & Speech*, 3, 7-10.

ONIONS, C. T. (ed) (1933). *The Shorter Oxford English Dictionary*, 2 vols., 1st Ed. Oxford: Oxford University Press.

SEASHORE, R. H. & ECKERSON, L. D. (1940). The measurement of individual differences in general English vocabularies. *J. educ. Psychol.* 31, 14-38.

SHANNON, C. E. (1948). A mathematical theory of communication. *Bell. Syst. Tech. J.*, 27, 379-423, 623-656.

SHANNON, C. E. (1951). Prediction and entropy of printed English. *Bell. Syst. Tech. J.* 30, 50-64.

SMITH, M. K. (1941). Measurement of the size of general English vocabulary through the elementary grades and high school. *Genet. Psychol. Monogr.* 24, 311-345.

TAYLOR, W. L. (1955). Applications of 'cloze', and entropy measures to the study of contextual constraints in samples of continuous prose. *Dissert. Abstracts*, 15, 464-465.

TREISMAN, A. M. (1962) Attention and Speech. D.Phil. Thesis, University of Oxford.

Manuscript received 8 June 1962

Acquiescence Response Set and the Questionnaire Measurement of Personality

By J. B. KNOWLES

*Medical Research Council's Clinical Psychiatry Research Unit,
Graylingwell Hospital, Chichester*

Data derived from normal subjects given three questionnaire measures of rigidity, and an independent measure of acquiescence response set are presented to demonstrate (1) that questionnaires of the 'true-false' type can be differentially prone to acquiescence response set, and (2) that the use of so-called 'balanced' scales as a means of counteracting acquiescent tendencies is of dubious value.

Studies have shown that when answering questionnaires and scales of the 'True-False', 'Agree-Disagree' type, subjects may consistently agree with items, or mark them as being true, although such items may express varying or even antagonistic psychological characteristics (e.g. Jackson & Messick, 1958, 1961). This response tendency—or set—to agree or acquiesce was first systematically described by Cronbach (1946). In his view acquiescent tendencies, in common with other response sets, 'influence performance most on ambiguous or relatively unstructured items' (p. 480); a view recently endorsed by Edwards (1961).

In the hands of other authors, however, the concept of acquiescence has been modified to the extent that it virtually denies the importance of item content. Thus, Couch and Keniston (1960) define it as 'a general tendency to agree or disagree with questionnaire items, *regardless of their content*' (p. 152, my italics). Although in their study overall agreement scores (obtained by scoring subjects' responses to a scale of 360 heterogeneous items solely for degree of agreement) had both a high reliability and close correlations with the number of times the subject responded 'true' to MMPI items and 'yes' to items of Cattell's 16 PF tests, of equal interest is the fact that with one exception agreement scores had virtually zero correlations with scores on Thurstone's Temperament Schedule, despite the fact that in six of the seven scales scores are comprised largely of 'yes' responses. Had the subjects consistently acquiesced, these correlations should have been high and positive rather than near zero.

A similar finding was apparent in a recent study by the present writer. Namely, scores designed to be a measure of acquiescence were found to be reliable, but to have relatively low correlations with scores on three questionnaire measures of rigidity, even though scores on these scales consisted largely, if not wholly, of the number of times the subject marked the item as being true. The first aim of the present paper is to present and discuss these data.

The second aim is to discuss the limitations of balanced scales as a means of controlling for acquiescent tendencies. Emphasizing the contribution of acquiescence to variance on personality and attitude scales, some authors have argued for the adoption of balanced scales, that is, scales in which 'true' and 'false' responses

contribute equally to a maximum scale score. 'This type of scale', state Couch and Keniston (*op. cit.*, p. 170), 'is a useful method of cancelling out at its source the contaminating influence of agreement response set on sum scores, and thus provides at least a statistical solution to some of the difficulties raised by this biassing tendency.'

However, if, as Couch and Keniston maintain, acquiescence is a general tendency to agree with questionnaire items *regardless of their content* then either the reversal of item *content*, or the selection of items which describe the absence of the trait in question, will alter the experimenters' scoring procedure (such items will now be keyed 'false') but not necessarily the subject's behaviour. As a consequence sum scores will continue to be biased. Similarly, if response sets are primarily elicited in situations where items are difficult or ambiguous, then reversal of the same items will leave the subject's behaviour unaffected, unless such reversals also reduce their ambiguity. On the basis of either argument balanced scales are likely to be as prone to bias by acquiescence as those scored in one direction. This expectation is confirmed by data presented in a later section of this paper.

METHOD

Tests

The data to be presented were collected during a study designed to assess the reliability and concurrent validity of three questionnaire measures of rigidity. The three scales were: (1) the rigidity scale described by Gough (1952), (2) the Nigniewitzky rigidity scale as published by Brengelmann (1960), and (3) a 38-item version of the Wesley Rigidity Inventory as published by Chown (1960).

An initial sample of 200 normal subjects were randomly divided into two equal groups. One sample received scales (with minor modifications in wording) in their published form; on these scales rigidity was indicated either predominantly or wholly by the subject marking the statement as being true. The remaining subjects received a balanced form of the same three scales; in these balanced scales up to 50 per cent of the original items were reversed, and the response 'false' taken to indicate rigidity on these items. In the main these reversals were achieved by relatively minor changes in wording, for example, by replacing 'dislike' for 'like', 'few' for 'most', 'rarely' for 'often', etc.

Subjects were also given the 70 items of the social introversion (Si) scale of the MMPI. The number of times the subjects responded 'true' to these items was taken as a measure of his degree of acquiescence. Since the Si scale is so constructed that a 'true' response contributes to a score on introversion on only 48 per cent of the items, high or low acquiescence scores cannot be produced by subjects consistently responding to the content of the items.

The 74 items making up the rigidity scales, and the 70 items of the MMPI Si scale were given in random order as one continuous questionnaire; each item could be answered as being either true or false.

Subjects

The total questionnaire was administered to normal subjects by a group of Psychiatric Social Workers. Of the 200 questionnaires sent out, 137 were returned. Seventy-nine of the questionnaires were returned by subjects who had received questionnaires containing the balanced scales. Apart from this difference in sample size, these subjects did not differ significantly in terms of mean age, sex and social class distribution from those who returned original scales. The total sample was predominantly female (66 per cent), middle-aged (Mean = 39.9; S.D. = 11.03), and of high social status (81 per cent being either social class I or II).

In summary, the design of this study provided a means of assessing the contribution of acquiescence to scores on each of the three rigidity scales in both their original and balanced forms.

RESULTS

Original scales

The correlations between scores on each of the three rigidity scales in their published form and the measure of acquiescence response set, together with their respective means, standard deviations and split-half reliabilities (corrected for length) are given in Table 1.

Table 1. *The intercorrelations between scores on rigidity and acquiescence and the (corrected) split-half reliabilities, means and standard deviations of these scales (N = 58)*

Scale	Number of items	Mean	S.D.	Corrected split-half reliability	Correlation with acquiescence
Acquiescence	70	29.48	6.45	0.732	—
Nigniewitzky Scale	14	4.19	2.41	0.408	0.408†
Gough Scale	22	9.33	3.75	0.799	0.295*
Wesley Scale	38	17.55	6.14	0.796	0.114

* $p = <0.05$; † $p = <0.01$.

Regarding acquiescence; the distribution of 'true' scores on the MMPI Si scale indicates that some subjects used 'true' and 'false' responses more than would be expected had they consistently responded to the designed content of the scale (i.e. responding 'true' to Introversion items and 'false' to Extraversion items, or vice versa) and the high reliability of these scores indicates that overall subjects were consistent in their use of these response categories. However, before it can be concluded that such behaviour is indicative of acquiescence it is necessary to exclude other interpretations of these data.

The first possibility is that high 'true' scores might have been produced by an inconsistent, but nonetheless valid, response to the content of the item. That is, subjects could respond 'true' to both introvert and extrovert items not because of acquiescence, but because such responses validly reflected 'inconsistencies' in their behaviour. However, whilst it is reasonable to expect some individuals to show such a response pattern, it is unlikely to account for the high split-half reliability of 'true' scores in the total sample. Further, in an extensive factor-analytic study using prison, hospital and college samples, Jackson and Messick (1961) found that although the Si scale was one of the few MMPI scales to possess validity, in that scores derived separately from 'true' and 'false' keyed items loaded on the same content factor, in all three samples these scores had high positive and negative loadings respectively on a factor of acquiescence. Thus, we may conclude that whilst a valid response to content cannot be excluded, there is clear evidence that acquiescence makes a significant contribution to variance on this scale.

A second possibility is that subjects responded 'true' or 'false' according to the perceived social desirability of the item. If the items of the Si scale preponderantly express either socially desirable or undesirable behaviour and personality characteristics, then subjects consistently responding according to their social desirability

would also consistently agree or disagree with these items; thus, their behaviour could be interpreted as reflecting either an acquiescence or social desirability response set.

Taking the proportion of subjects responding 'true' to an item as an index of its social desirability, the distribution of item desirability value was found to be approximately normal. Since, by definition, either a large or small proportion of subjects indicate desirable or undesirable items to be true, 'true' scores derived solely from the items at the tails of the distribution are likely to act as a constant and, thus, there should be a high correlation between 'true' scores obtained from the total scale, and those obtained from items of neutral or near neutral social desirability. To test this notion items answered 'true' by either less than one-third or more than two-thirds of the sample were excluded, and a 'true' score on the remaining 36 'neutral' items correlated with the scores previously obtained from the total scale. The part-whole correlation was 0.92. Thus, we may conclude that although social desirability is likely to have affected responses to some items of the Si scale, such behaviour does not materially affect the interpretation of acquiescence scores as described in this paper.

Thus, data obtained from the Si scale provide evidence of acquiescence, that is, subjects consistently endorsed or rejected the items making up the scale. However, the intercorrelations between acquiescence and scores on the three rigidity scales indicate that a knowledge of acquiescent tendencies in one scale—the Si scale—enables us to predict little or no variance on other scales which, by design (all or the majority of items being keyed 'true' to indicate the scale direction), might be expected to elicit acquiescence. Although all three intercorrelations between acquiescence and rigidity scores were positive, in the case of the Wesley scale the correlation did not differ significantly from zero, and at best the intercorrelation was no higher than 0.408.*

Balanced scales

In calculating the results for the subjects who received the balanced form of the three rigidity scales, acquiescence scores were correlated with both the total rigidity score on each scale, and a true-false discrepancy score. This discrepancy score was the subject's rigidity score derived from items keyed 'true' minus his rigidity score derived from items keyed 'false' (20 being added to all values to eliminate negative signs). These correlations, together with the (corrected) split-half reliabilities of the total scales (each half having an equal number of 'true' and 'false' items), are given in Table 2.

As in the case of the subjects given the rigidity scales in their published form there was reliable evidence that subjects consistently responded 'true' (or 'false') to the items of the MMPI Si scale. The significant positive correlations between these

* It might be objected that these correlations should be corrected for attenuation. However, as McGee (1962) has pointed out, the very fact that scales which might be expected to measure acquiescence do *not* have high split-half reliabilities (e.g. the Nig涅witzky Scale in the present case) indicates that acquiescence is not an all-pervasive response set. To correct for attenuation, and thus assume perfectly reliable measures, would be to beg the present question.

Table 2. *The intercorrelations between scores on 'balanced' rigidity scales and acquiescence and the (corrected) split-half reliabilities of these scales (N = 79)*

Scale	Corrected split-half reliability	Correlation between acquiescence and total score	Correlation between acquiescence and discrepancy score
Acquiescence	0.776	—	—
Nignewitzky Scale	0.483	0.144	0.239*
Gough Scale	0.667	0.067	0.506†
Wesley Scale	0.807	-0.019	0.380†

* $p = <0.05$; † $p = <0.01$.

acquiescence scores and the true-false discrepancy measures also indicates that there was a tendency for the individual high on acquiescence to derive his rigidity score mainly from those items of the scale which were keyed 'true', and for the individual low on acquiescence to derive his rigidity score mainly from items keyed 'false'. These data do not support Couch and Keniston's view that 'balanced' scales should serve to cancel out any contamination due to agreeing response set' (*op. cit.*, p. 156).

In view of the above correlations it is not surprising that the correlations between acquiescence and total rigidity scores were virtually zero in all three cases. On any balanced scale, a subject who consistently responds 'true' (or 'false') will score at the mid-point of the scale when the items are scored for content; conversely the individual who consistently responds to content will score at the mid-point on acquiescence. Therefore, a correlation between acquiescence and content scores is likely to be zero in the total sample. Clearly, however, it would be erroneous to infer that this zero correlation indicates that scores are not biased by response set.

DISCUSSION

It is difficult to reconcile the first finding of this study, namely that scales alike in that their items were keyed 'true' were differentially prone to acquiescence, with the view that some subjects respond to the form of questionnaire items regardless of their content. Whilst this hypothesis might be able to account for the consistent 'true' responding on the MMPI Si scale, it cannot account for the generally low correlations between acquiescence and rigidity scores.

A similar conclusion is reached in a recent study by McGee (1962). He gave six putative measures of acquiescence incorporating both structured and unstructured items and found first, that two-thirds of his measures had reliabilities of 0.57 or below, and second, that although all but one of his intercorrelations were positive, all but three were below 0.3. He concluded that the most economical interpretation of his data was that 'There is no general trait of response acquiescence which appears reliably in different situations' (p. 232).

In view of these findings it may be of interest to consider why some scales should be vulnerable, whilst others are not. Eysenck (1962) has suggested that social attitude scales may be particularly prone to bias by acquiescence because their items express general principles about which the responder may have no clear or strong views. In contradistinction, personality scale items are likely to arouse

interest because they are primarily self-descriptive, and one may answer them without fear of contradiction. A similar viewpoint is expressed by Chapman & Campbell (1959) who found no evidence of acquiescence in the answers to items of the Taylor Manifest Anxiety Scale and suggested that 'this is probably attributable to the quite specific and personal reference of the items'. However, this argument cannot account for the fact that acquiescence has been shown to contribute significantly to variance on such personality scales as the MMPI and Cattell's 16 PF test (Couch & Keniston, 1960; Jackson & Messick, 1961).

A more likely factor is item 'difficulty' or ambiguity. As noted previously Cronbach (1946) and Edwards (1961) suggest that acquiescence, in common with other response sets, is elicited in ambiguous or unstructured situations, that is, when the subject experiences some doubt as to the appropriate or correct response. Reading through the scales in the present study one is struck by the florid and prolix style of the Nigniewitzky items, scores on which had the highest correlation with acquiescence, whereas the items of the Wesley scale, which had a zero correlation with acquiescence, appear to be somewhat briefer descriptions of specific behaviour.

The reversal of an item, or the selection of items describing the absence of the trait in question, is likely to increase the ambiguity of a scale since responses to these items may now involve a double negative. On this basis one might anticipate that balanced scales would be particularly vulnerable to acquiescent tendencies, and this might well account for the fact that acquiescence bore no relation to scores on the Wesley scale in its original form, but biased scores when one-half the items were reversed.

The use of balanced scales as a means of counteracting the bias arising from acquiescence would seem to be of dubious value for two other reasons. First, as noted earlier, whether acquiescence is defined as a set to respond to questionnaire items in a particular way irrespective of their content, or whether it is regarded as a tendency elicited by ambiguous content, modification of item *content* will not alter the behaviour of the acquiescent individual. The high split-half reliability of 'true' scores on the Si scale, itself a balanced scale, and the significant correlations between acquiescence and the true-false discrepancy scores, provide empirical support for this viewpoint. Secondly, the avowed purpose of a balanced scale is nullified if the scale when scored unidirectionally does not elicit acquiescent tendencies. The data of this and other studies quoted suggests that this may not be a rare occurrence, and it would seem valuable for future research to concentrate as much on defining the characteristics of scales which do *not* elicit acquiescence, as upon those that do.

I am grateful to members of the Parents Group of the Association of Psychiatric Social Workers for collecting the data on which this study is based.

REFERENCES

BRENGELMANN, J. C. (1960). Extreme response set, drive level and abnormality in questionnaire rigidity. *J. ment. Sci.*, **106**, 171-186.

CHAPMAN, L. J. & CAMPBELL, D. T. (1959). Absence of acquiescence response set in the Taylor Manifest Anxiety Scale. *J. consult. Psychol.*, **23**, 465-466.

CHOWN, Sheila, M. (1960). A factor analysis of the Wesley Rigidity Inventory: its relation to age and non-verbal intelligence. *J. abnorm. soc. Psychol.* **61**, 491-494.

COUCH, A. & KENISTON, K. (1960). Yeasayers and Naysayers: agreeing response set as a personality variable. *J. abnorm. soc. Psychol.* **60**, 151-174.

CRONBACH, L. J. (1946). Response sets and test validity. *Educ. psychol. Meas.* **6**, 475-494.

EDWARDS, A. L. (1961). Social desirability or acquiescence in the MMPI? A case study with the SD Scale. *J. abnorm. soc. Psychol.* **63**, 351-359.

EYSENCK, H. J. (1962). Response set, authoritarianism and personality questionnaires. *Brit. J. soc. clin. Psychol.* **1**, 20-24.

GOUGH, H. (1952). The Development of a Rigidity Scale. Mimeo. Institute of Personality Assessment and Research, University of California, Berkeley.

JACKSON, D. N. & MESSICK, S. (1958). Content and style in personality assessment. *Psychol. Bull.* **55**, 143-252.

JACKSON, D. N. & MESSICK, S. (1961). Response styles and the assessment of Psychopathology. In *Measurement in Personality and Cognition* (ed. S. Messick and J. Ross). New York: Wiley.

MCGEE, R. K. (1962). The relationship between response style and personality variables: I. The measurement of response acquiescence. *J. abnorm. soc. Psychol.* **64**, 229-233.

Manuscript received 29 June 1962

Social Differences among Day-release Students in Relation to Their Recruitment and Examination Success

BY ETHEL C. VENABLES, M.Sc., PH.D.

Nuffield Research Fellow, Institute of Education, University of Birmingham

It is argued in this paper that failure and wastage in part-time day classes is not primarily a matter of lack of ability, but can only be understood in relation to their heterogenous nature and the differential motivation which brings together this diverse group of young people. Youths whose fathers are in unskilled or semi-skilled jobs fared at least as well, and often better than, those from professional families. This should not be interpreted as meaning that the family ceases to exert influence once a young man goes to work. Rather, it is one of several indications that those male employees under 18 who attend first year classes on one day per week are a skewed sample of the population. Most of them come from an artisan background and their success rate on the first year of just over 50 per cent appears to depend at least as much on their individual motivation as on their ability as measured by test scores. At the two extremes the groups tended to be unrepresentative; those from the professional and managerial backgrounds being mostly the intellectually inferior members and those from Social Classes IV and V being the more intelligent members of those groups.

I. INTRODUCTION

The Barlow Committee report on Scientific Manpower, published in 1936, recommended that the Universities should double their output of scientists in ten years. They also maintained that 'even if the total student population in British Universities were doubled, this country would still fall far short of a number of European countries and the United States of America in the relative provision which it makes for higher education'.

From then on discussion on University expansion, selection and manpower problems filled the educational press. After one such consideration of the supply of first rate ability, the Editor of the Universities Quarterly (1950) commented 'we have little reliable knowledge on which to base an answer to these questions'. Later (1952) he summed up the issues in these terms '... any access of strength must surely come from among the very large number of young people who are at present receiving part-time training at technical colleges'.

II. PROCEDURE

As part of the attempt to answer some of these questions a survey of the ability of young technical college students was begun in 1950. Three further investigations have been carried out since then in 1952, 1957 and 1960, and in each of them verbal and non-verbal tests of intelligence have been given to first year students in day-release courses. In addition, information about father's job, family size and previous schooling has been collected in order to examine to what extent the various social groups avail themselves of this opportunity for further education. Thus the results fall into three sections:

- Section 1. What is the spread of ability among part-time day-release students in technical college classes?
- Section 2. How far are they representative of school leavers from the various types of secondary school and of the various social groups in terms of their fathers' occupation?

Section 3. How far is success on these courses related to previous schooling and family background?

In assessing father's job the five point scale of the Registrar General's classification was used: Classes I and II covering professional and managerial employment, Class III—clerical and skilled manual workers, and Classes IV and V—semi-skilled and unskilled labourers. Seven local technical colleges* in two urban areas have co-operated.

In most cases the first year of a part-time course in a technical college is designed for the school leaver who has not reached G.C.E. 'O' level. Entrants who have passed in four 'O' level subjects including Mathematics and a Science are usually exempt from the first year. In planning the investigations it was assumed that these 'second year entrants' were able young people who could be considered to belong to the 'pool of ability' so much under discussion. The Central Advisory Council (1954) which investigated 'Early Leaving' found that, of the boys and girls entering the Grammar Schools in 1946, a third of those considered by the schools to be capable of taking an advanced sixth form course left at sixteen. It is a safe assumption that many of these, having succeeded at 'O' level, took up industrial apprenticeships and found their way into the second year of a technical college course.

Details of the results under the first heading have been reported elsewhere (Venable, 1960, 1961a) and summaries only are presented here. A succinct account of the organization of these first year classes is to be found in Chapter 33 of the Crowther Report *Fifteen to Eighteen* entitled 'Beginning Further Education', Central Advisory Council (1959). In dealing with the problems raised in placing students in suitable courses these results are reported and their implications discussed.

Part-time day release. The idea of releasing young people in industry for one day per week to attend technical college classes dates from 1918 when the Fisher Education Act required authorities to establish continuation schools for young people between the school-leaving age of 14 and the age of 16, and enforce their attendance for 320 hours per annum. The story of this early attempt at further education for the early school leavers and its disappointing failure has been told by P. I. Kitchen (1944) Principal of the Technical College at Rugby, the only local authority to carry out, continuously since 1919, the intentions of the Act. After the Second World War the 1944 Education Act raised the school-leaving age to 15 and again made provision for continued education in county colleges—part-time if need be—this time to the age of 17. Again, this part of the Act has not been implemented; very few county colleges, as such, have been established, and nowhere, except still at Rugby, is release compulsory.

Since 1944 there has, nevertheless, been a steady increase in the number of students between 15 and 17 and up to 21 and beyond, released voluntarily by firms for attendance at technical colleges on one day per week, and the figure is now about 19 per cent of those of the 15-17 age group who are in employment.

It is important in interpreting the figures to note that day release among girls is small—about one-fifth of the total. If boys only are considered, the proportion of those 'at risk' who in fact attend part-time day-release classes is about 30 per cent. Thus, though 43 years have passed since the Fisher Act, in which 'day release' was first recommended, 70 per cent of young male employees still get no further education after leaving school.

Recruitment. The traditional role of the technical college has been to provide a second chance for those who achieve no qualifications at school and a route to professional employment for those who cannot (or do not wish to) continue in full-time education. Over half the Associate members of the Institution of Mechanical Engineers, which has a membership of over 40,000, have achieved their qualification (A.M.I.Mech.E.) in this way.

The students, particularly those in engineering departments, can be roughly divided into two groups: those taking National Certificate courses (which can lead on to the professional qualifications already mentioned) and those taking City and Guilds trades and operative courses. The former tend to be somewhat superior on intelligence and other tests, as various *selection sieves*—including self-selection—operate, to prevent the poorer candidates attempting the more difficult courses.

* Since publication of the White Paper on Technical Education by H.M.S.O. in 1956, the technical colleges of England and Wales, numbering nearly 500, have been reorganized into four groups—Local colleges, about 300; Area Colleges, about 150; Regional Colleges, 25; and Colleges of Advanced Technology, 10.

Since the publication of the White Paper 'Better Opportunities in Technical Education' in January, 1961, the pattern of courses is changing, but the essential problems remain: how to increase the proportion of young people continuing their education after leaving school and how to detect the able ones among them and encourage more of them to gain higher qualifications.

Method. The following tests have been used:

1. Vernon's Arithmetic/Mathematics test.
2. Raven's Progressive Matrices (1947).
3. An accuracy score devised from 2.
4. NIIP Form Relations test.
5. Mill Hill Vocabulary Scale.
6. NIIP Group Test 33.
7. A.H.4 Test of General Intelligence, Parts I and II.
8. Maudsley Personality Inventory (MPI)—Scales 'N' (Neuroticism) and 'E' (Extraversion).

Success in the first year examination has been assessed in three ways:

- (i) Pass rates;
- (ii) Correlations of standardized examination marks with test scores;
- (iii) Achievement Ratios.

The first two alone are inadequate as they are not related to the spread of ability in the class or to the level of ability needed for success. The Achievement Ratio is an attempt to determine how far achievement matches ability. It is a measure of the pass rate with test results held constant.

The four tests giving maximum prediction were first isolated by a multiple regression analysis. The sum of these four standardized scores was calculated for each student and frequency tables drawn up according to pass/fail on each course. By the method of least misfits a 'cut-off' score for each examination was calculated. This yields four groups of students:

1. Those above 'cut-off' who pass—Expected Pass (E.P.).
2. Those above 'cut-off' who fail—Under-Achievers (U.A.).
3. Those below 'cut-off' who fail—Expected Fail (E.F.).
4. Those below 'cut-off' who pass—Over Achievers (O.A.).

An Achievement Ratio. Achievement can now be measured by the ratio of those above the 'cut-off' who passed as expected, over the total who it is assumed were capable of passing,

E.P.

i.e. $\frac{\text{E.P.}}{\text{E.P.} + \text{U.A.}}$

The basic assumption underlying any selection procedure is that each subject wishes to succeed in the job or course of study for which he is being selected. This assumption cannot always be made with confidence in the case of these students, since some of them come only because the firms insist on their attendance, and in some industrial situations there is no immediate or obvious gain arising from successful performance.

It is suggested that the achievement ratio might be regarded as an indication of the general level of motivation within the groups.

III. RESULTS

Section I

(a) *Level of ability—verbal and non-verbal test results*

In the testing programmes carried out in 1950, 1952 and 1957 the following two tests (among others) were used:

1. NIIP Group Test 33—a test of verbal ability.
2. Raven's Progressive Matrices 1947—a non-verbal test.

The results showed no statistical differences between the colleges or the geographical areas, and the level of achievement on non-verbal tests was, in each case,

very much higher than on verbal ones. The combined results for first year Ordinary National Certificate students are shown in Table 1.

Table 1. *Percentages of first year part-time day-release engineering students scoring within the same range as the top half of the population on two different types of intelligence tests. Ordinary National Certificate students only*

General population	N > 1,000	
	Part-time engineering students	
	Raven's Matrices	NIIP G.T.33
	1947	
Top 10%	37%	7%
Top 20%	60%	16%
Top 30%	77%	27%
Top 50%	87%	58%

This latter result was not, of course, surprising, since the relative lack of verbal facility on the part of such students has not gone unnoticed either by technical college teachers or employers. The extent of the difference was, however, unexpected; 40 per cent of the first year Ordinary National Certificate students had scores at or above those obtained by 75 per cent of University students on the 1947 version of Raven's Matrices, which was designed for those of superior ability. On a verbal test (NIIP Group Test 33) not more than 6 per cent were comparable in this way with University students (Venables, 1960, 1961a).

Thus any assessment of the level—or reserve—of ability among these part-time day-release students will depend upon the kind of measure used. Most of them are studying engineering, and even at University level, engineering students are usually superior to other University students on non-verbal tests and less able (or less interested) in verbal ones.

(b) *Improvements in verbal ability during three years of college study*

Examination results for the 1957 sample were recorded up to 1960, and a few months before their third year examinations, just under 200 of them worked the same tests again. The improvements in verbal test scores obtained indicate continued intellectual development after the age of 16 years. A further sample of third year students was tested for the first time to enable an allowance for practice effects to be made. Table 2 summarizes these results which have been fully reported elsewhere (Venables, 1961b).

Table 2. *Changes in percentile ranking in verbal and non-verbal intelligence tests on retesting after an interval of 2½ years. (Retest results reduced to allow for practice effects.) O.N.C. students only*

Norms for the general population	G.T.33 Verbal		Raven's Matrices	
	1957	1960	1957	1960
Above 75th percentile	18.6%	33.1%	71.4%	83.4%
50th-75th percentile	35.2%	42.0%	18.7%	11.9%
Below 50th percentile	46.2%	24.9%	9.9%	4.7%

Section 2

(a) *Recruitment according to previous schooling and father's occupation grade*

Three samples—1952, 1957 and 1960—have been analysed under these two heads with very similar results. The 1960 figures are given in Table 3. Almost all the ex-Grammar school pupils who enter first year technical college classes attempt the Ordinary National Certificate courses. Ex-secondary Technical school pupils are over-represented in both types of classes.

Table 3. *Recruitment analysed according to previous schooling and social class*(a) *According to previous schooling*

	All O.N.C.s 1960 N = 195	Trade 1960 N = 380	Proportions in school population. Ministry of Education Report 1959/60
Grammar	18%	3%	15%
Technical	23%	6%	5%
Modern*	59%	91%	80%

* The few students from comprehensive and special schools included here.

(b) *According to father's occupational grade*

	All O.N.C.s 1960 N = 195	Trade 1960 N = 380	Census 1951
I and II	18%	11%	18%
III	70%	65%	53%
IV and V	8%	17%	29%
Retired, deceased and not known	4%	7%	—

Professional and managerial families in this sample are found to contribute a representative proportion even to first year classes, but the semi-skilled, and particularly the unskilled homes, are very much under-represented. About three-quarters of part-time technical college students come from Social Class III. Something like a tenth of their fathers would be classified as non-manual workers: thus in taking up apprenticeships a high proportion are following in father's footsteps.

(b) *Test scores according to previous schooling and father's occupation grade*

Results from all samples tested have been very similar, and those summarized in Table 4 are again from the most recent sample (1960). They refer to the whole of the first year entry of students into one college in the Midlands. About a fifth were in Ordinary National Certificate (O.N.C.) classes, and the rest were divided between 12 City and Guilds courses.

The intelligence and attainment tests follow the expected pattern in relation to previous schooling, and most of the differences between the three main types of school are significant. When analysed by social class, differences are mostly insignificant, and often not in the expected direction. The few from Social Class I

Table 4. First year students in all courses in one local technical college. Session 1960/61. N = 575

Mean test scores according to previous schooling							
Previous schooling	Per cent sample	Vernon's Arith/Maths	Raven's Matrices 1947	Mill Hill Vocab.	NIIP Form Relations	MPI 'N' Scale	MPI 'E' Scale
Grammar	8	52.4	30.1	48.4	40.4	22.2	29.8
Technical	11	51.3	28.8	44.1	37.1	24.1	84.1
Modern	75	42.7	26.0	40.8	33.5	23.0	82.3
Other, including Comprehensive	6	45.1	25.5	43.1	36.4	21.7	70.0
Totals	100	44.6	26.2	41.8	34.5	23.0	76.4
						28.7	72.8

Mean test results according to social class							
Social Class	Per cent sample	Vernon's Arith/Maths	Raven's Matrices 1947	Mill Hill Vocab.	Form Relations	MPI 'N' Scale	MPI 'E' Scale
I	6	47.3	25.6	43.8	34.8	25.6	29.5
II	8	46.0	26.8	41.8	35.9	22.7	70.5
III	66	44.5	26.4	42.2	35.3	22.9	31.0
IV, V	14	42.0	25.2	40.2	32.3	20.8	27.9
Not known	6					29.1	74.1
						69.3	69.3

Population Norms	Maths age 16 Score = 50	25 is score at 70th decile	42 is score at 50th decile	33 is score at 50th decile	Mean Score 19.89	Mean Score 24.91	69 is score at 90th decile
Means for University students		32.9	62.67	35.0*	25.17	22.67	96.36

* Mean for Secondary Grammar School boys aged 16 and over.

This is a simplified table omitting details of numbers and s.d.s. The significance of the differences is dealt with in the text.

have the best attainment scores (Vocabulary and Maths) but are inferior on the two intelligence tests (Raven's Matrices and A.H. 4). They also have higher scores on the Neuroticism scale of the Maudsley Personality Inventory. On this scale the mean score for the sample was greater than for the population as a whole but less than the mean for University students. The 'E' scale results showed them to be more extravert than the general population, while the University student norm is lower.

The group as a whole, i.e. including all those in City and Guilds classes, ranks in the top half of the population on most of the tests, and on the Matrices and A.H. 4 they are in the top quartile.

(c) *Verbal and non-verbal differences according to previous schooling and fathers' occupational grade*

To make a comparison between the verbal Group Test 33 and the non-verbal Matrices involves the use of standard scores. By using the A.H. 4 test of general intelligence (as was done with the 1960 sample) a direct comparison between the two types of test could be made as this test is in two parts: Part I being verbal (and numerical) and Part II—non-verbal. It also provided a check on the previous results. Verbal and non-verbal differences with the two different sets of tests are analysed in Tables 5 and 6 showing that the pattern of scores is the same in both samples.

On the verbal tests mean differences according to previous schooling are in the expected direction—significantly so in the case of the larger group. Non-verbal differences between Grammar and Modern school pupils were negligible, but the means of the ex-Technical school pupils were superior to each of the others. In the case of the 1957 sample, again the numbers involved were large enough for the differences to be statistically significant.

It has been demonstrated many times that scores on intelligence tests for the population as a whole decrease as we descend the social scale. When the results for these students are analysed in this way the skewed nature of the sample becomes very apparent.

Social differences between mean scores on the Matrices test were negligible, and in the 1960 sample non-verbal mean scores actually increase, though not significantly, as we descend the social scale. As judged by this kind of test it is the less intelligent sons of professional men who attend such classes, and the few from Classes IV and V who continue to study in this way tend to be above the average for their group. The two highest occupational groups have higher verbal test scores, but those from Classes IV and V are not inferior to those from Class III. The influence of family background is most plainly seen in the differences between the two types of score; the sons from Classes I and II showing the least discrepancy and those from IV and V the greatest.

(d) *Family size*

A further indication that the population of part-time day-release students is an unrepresentative cross-section of the population as a whole is provided by a study of their intelligence in relation to size of family.

Table 5. *Mean scores on verbal and non-verbal tests according to previous schooling. O.N.C. students only*

	1957 sample. N = 384						1960 sample (Mech.Eng. only). N = 118					
	NIIP G.T. 33			Raven's Matrices 1947			A.H. 4.I			A.H. 4.II		
	Per cent of group	M	S.D.	M	S.D.	Per cent of group	M	S.D.	M	S.D.	M	S.D.
Previous schooling												
Grammar	24	119.25	22.2	26.87	5.65	16	39.34	6.5	44.68	10.05		
Technical	24	105.37	15.2	29.24	6.17	23	38.56	6.2	46.00	7.56		
Modern	49	98.08	17.3	26.61	5.68	57	35.14	6.6	43.96	8.28		
Comprehensive and others	3	—	—	—	—	4	—	—	—	—	—	—
Means for University students		156.31		32.90			48.0		49.0			
									approx.			

Table 6. *Mean scores on verbal and non-verbal tests according to father's occupational grade. O.N.C. students only*

Father's occupational grade	Per cent of sample	1957 Sample. N = 384				1960 Sample (Mech. Eng. only). N = 118			
		NIIP G.T. 33		Raven's Matrices 1947		A.H. 4.I		A.H. 4.II	
		M	S.D.	M	S.D.	Per cent of sample	M	S.D.	M
I and II	14	112.35	19.6	28.1	7.1	21	38.07	6.9	44.32
III	61	104.6	20.2	27.7	6.0	68	36.37	7.1	45.01
IV and V	6	105.1	17.2	28.4	4.9	8	36.70	7.5	46.60
Not known	19	104.51	18.0	28.4	4.9	3	40.00	4.2	45.01
Totals	100	105.59	19.70	27.93	5.9	100	36.85	7.0	45.01

The Scottish Mental Survey found a negative correlation between intelligence scores and family size (Scottish Council for Research in Education, 1949). To what extent such an effect is genetic or environmental has been discussed by Nisbet (1953) and it seems very probable that part at least of the correlation is due to heredity. Using scores on a verbal test for over 1000 of these first year part-time day-release students no significant correlation of this kind was found (Warburton & Venables, 1956).

Assuming the validity of the negative relationship, this result would be achieved if the more intelligent members of the larger families chose this form of further education while their peers from smaller family units were able to attend full time courses.

Section 3

(a) Pass rates

Fifty-six per cent of the 1957 sample ($N = 384$) were successful in the first year examination. For the social groups analysed in Tables 4 and 5 the results were as follows:

Previous schooling	1957 sample	
	(per cent)	
Grammar	62	
Technical	61	
Modern	52	
<i>Social Classes</i>		
I and II	55	
III	57	
IV and V	54	
Totals	56	

These pass rates show only small differences according to family background and are similar to those found in the Technical Courses Survey conducted at the request of the Crowther Committee. In Part II of their Report *Fifteen to Eighteen* they draw the somewhat hasty conclusion that this similarity between the social classes supports the 'view that home background has more effect at the school stage than during further education . . . little further effect of parental occupation can be seen once the student has entered employment and begun a part-time course' (Central Advisory Council, 1959). Further analysis of the 1957 cohort suggests that the skewed nature of the population is the more likely explanation. After three years of study differences were more marked, and scores on the Raven's Matrices continued to be more predictive of success (Table 7). Scores on the verbal test for successful students differed significantly between the social groups and bore little or no relationship to success for the sample as a whole. The 9 per cent from the professional and managerial classes who succeeded in gaining a certificate in minimum time had a mean verbal score in the top decile, whereas students in Social Class 3 who succeeded had a mean score at the 50th percentile. A small

Table 7. *Percentages gaining Ordinary National Certificate in minimum time, according to father's occupational grade*

Occupational Grade	Per cent of group	Mean test scores on entry of those obtaining the certificate after 3 years of study					
		Group test 33			Raven's Matrices 1947		
		M	S.D.	M	S.D.	M	S.D.
I and II	14	{	Enter	100	52	22	135.15
		{	Qualify	55	24	9	22.48
III	61	{	Enter	100	49	23	110.51
		{	Qualify	57	25	8	21.13
IV and V	6	{	Enter	100	46	23	120.75
		{	Qualify	54	23	15	12.65
Father's job not recorded	16	{	Enter	100	38	12	114.75
		{	Qualify	47	12	5	17.09
Father deceased	3	{	Enter	100	45	27	107.39
		{	Qualify	45	36	27	13.60
Totals	100 N = 384	{	Enter	100	49	20	115.57
		{	Qualify	56	23	9	21.51

group of students who reported 'no father' or 'father deceased' appeared to be highly motivated. Three of the 11 (27 per cent) obtained a certificate in the minimum time with verbal test scores on entry lower than for any other successful group. By contrast, of the 60 students who answered the questionnaire with least care and interest and had, among other omissions, failed to state their father's job, correlations with success on the first year examination were negative and only three (5 per cent) gained a certificate in three years.

(b) *Achievement in relation to test scores*

The matrices of inter-test correlations have been very similar for all the research programmes. Correlations between the five cognitive tests and examination results for the 1960 sample of O.N.C. students only, were as shown in Table 8.

Table 8

	Maths	Raven's Matrices	Mill Hill Vocab.	Form Relations	A.H. 4
Product moment r between tests and total standardized exami- nation marks	0.408	0.282	-0.015	0.201	0.173

In all four testing programmes the Maths test has invariably proved to be the best single predictor of success in the first year examination, and the verbal tests the least useful. For example, in the 1957 sample there was a difference of less than five points on Group Test 33 between the mean scores of those who passed and those who failed, the actual figures being 107.70 and 102.75. For the grammar school group the comparable figures were 119.25 and 117.87, and for the ex-modern school pupils 98.08 and 96.91. Thus the latter group succeeded with a mean score very much less than that of the Grammar school failures.

(c) *Achievement in relation to previous schooling and family background*

The four tests which together gave the best prediction of examination success were: Vernon's Arithmetic/Mathematics test, Raven's Progressive Matrices (1947), NIIP Form Relations and A.H. 4, and the use of these in a combined standardized score to determine the right 'cut-off' for each course separated the 480 students in the 1960 sample in the following way:

	Per cent
Expected Pass	65
Under Achievers	15
Expected Fail	14
Over Achievers	6

This gives an overall achievement ratio of 81 per cent, i.e. 19 per cent of those with scores above the 'cut-off' failed. The results shown in Table 9 were obtained for the various social groups. Where the ratios differ from the 'expected' figure, the level of probability that the difference is not due to chance is indicated.

Table 9

<i>Social Grouping</i>	<i>Achievement ratios</i>
<i>Previous schooling</i>	
Grammar	69%
Technical	81%
Modern	81%
<i>Social Class</i>	
I and II	62%
III	80%
IV and V	86%
Firms with 1000 and more employees	88%
Firms with under 1000 employees	69%
Course of study	Varied from 100% to 54%

The results suggest that social factors have some bearing on under-achievement, and in so far as this is indicative of poor motivation it would seem that the secondary modern school leaver and the child of the semi- and unskilled manual worker tends to take his chance of further education at the local technical college more seriously than his classmate from the Grammar school or from a superior social class. This confirms the extended results for the 1957 sample given in Table 7.

Further, it might be postulated that the large firms are more interested to encourage their apprentices to achieve qualifications.

The course in which the achievement ratio was 100 per cent served a trade where there are many fewer jobs than applicants, and the students are correspondingly enthusiastic. On another which had an achievement ratio of 96 per cent, the industrialists take the keenest possible interest, and there is general acceptance by employees and employers alike of the necessity for academic study. Some students in each course manage to reach pass level by the end of the year despite very low measured ability at the outset.

This over-achievement has not been found to be linked with social factors. The group is very small and a few of them were succeeding at the second attempt, but among those who could be regarded as 'genuine' over-achievers, personality factors—e.g. introversion and anxiety—appear to be the most important variables when they are compared with those with equivalent test scores who fail.

IV. DISCUSSION

The relationship between intelligence test results and the family's socio-economic and educational status has been recognized since the first large scale group testing carried out with the Army Alpha test during the First World War. Such findings have since been well documented, but discussion over the relative contributions of heredity and environment continues.

Sociologists naturally tend to lay stress on environmental differences, whereas much psychological work has in the past been directed towards the measurement, as far as possible, of pure inborn ability. The developmental approach of psychologists such as Hebb (1945) and Piaget (1950) has, however, underlined the

difficulty of producing 'culture-free' measures of this kind. They have shown that environmental factors influence intellectual development from the earliest years and that the effect tends to be cumulative. In so far as such intellectual development can be regarded as a process of conditioning, then temperamental differences in susceptibility to this process become important and reintroduce the question of the deterministic effects of heredity (Eysenck, 1959).

The situation is further complicated by the fact that attitudes to test situations and to particular varieties of tests are also likely to be culturally induced. For instance, the positive and negative discrepancies between verbal and non-verbal test scores shown by verbally fluent academics on the one hand and scarcely literate engineering apprentices on the other, may well owe something to differential motivation towards different varieties of test material in addition to any effects of temperamental, i.e. genetic differences.

The understanding of such cultural and class differences is a matter of great importance if the educational and industrial problems of a socially mobile welfare state are to be solved. More and more young people are crossing the educational and social barriers, and recent autobiographical writings by Richard Hoggart and others have drawn attention to some of the difficulties inherent at least in the Anglo-Saxon scene. Here, differences in speech and in the use of words and in attitudes to verbal skills are the most characteristic indicators of social differences. Many of the engineering apprentices, whose results are reported here, were as consciously aware of this as Hoggart (1957) or Martin Green (1961), though verbal expression of it was hardly so felicitous. 'When you're a working class lad like me and you come to a college like this it takes a bit o'time before you understand what folks are saying. You've nearly got to learn another language.'

In the chapter on 'Educational and Vocational Implications of Intelligence Tests' in his recent book, Vernon (1960) suggests that 'it is particularly unwise to use non-verbal group tests as a criterion of educational potentiality, except possibly among deaf or non-English-speaking pupils'. Results obtained with these young engineers, however, suggest that perhaps foreign pupils with minimal English and at least the partially deaf are not so much separate categories but simply at the extreme end of a continuum of verbal skill which might embrace next the bilingual group, followed by normal English-speaking peoples ranging from the semi-literacy of the unskilled labourer to the highly sophisticated fluency of the scholar.

At one of the colleges co-operating in these researches a small group of deaf students who have found employment in engineering firms are taught. The 1959 group had a mean score on Raven's Matrices (1947) at the 75th percentile, but they were only just able to attempt the first few items of the vocabulary and verbal tests. The work of Jones and Stewart (1951) with Welsh children has shown that bilingual children have significantly lower scores on a verbal test than monoglot English-speaking children, even when non-verbal test scores are held constant. Cronbach (1949) has suggested that 'it is possible that non-verbal tests deserve wider application in industry than they have received', and these results certainly indicate that with this type of student non-verbal test scores often give a better indication of potential success than verbal ones.

The five point scale of the Registrar General's is, of course, a very crude tool for the measurement of social differences. Certainly when considering the social attitudes affecting achievement it is of very limited use, and in a skewed population such as this one could be misleading. Those who at one extreme have failed to conform to their expected social role, and, at the other, those who hope to achieve by educational or other means, social advancement, are likely to display attitudes to education and success in direct contradiction to that expected from their social background. It has been found, for example, in a preliminary questionnaire relating to attitudes to practical and theoretical work that those in Social Classes IV and V showed on average a greater interest in theoretical work and a preference for using their brains, while those whose fathers were in Social Classes I and II were, on the whole, in favour of working with their hands.

A survey of the attitudes of these students to work and success has been made by Mr David J. Lee of the Nuffield Research Unit: how far these attitudes are influenced by family and place of work is a matter for further study.

REFERENCES

BARLOW COMMITTEE REPORT (1946). *Scientific Manpower*. H.M.S.O.

BURT, SIR CYRIL (1943). Ability and income. *Brit. J. educ. Psychol.* 13, 83-98.

CENTRAL ADVISORY COUNCIL FOR EDUCATION (ENGLAND) (1954). *Early Leaving*. H.M.S.O.

CENTRAL ADVISORY COUNCIL FOR EDUCATION (ENGLAND) (1959). *Fifteen to Eighteen*, Part II, 201. H.M.S.O.

CRONBACH, L. J. (1949). *Essentials of Psychological Testing*. N.Y.: Harper and Bros. *Education Act* (1944). H.M.S.O.

EYSENCK, H. J. (1959). Personality and verbal conditioning. *Psychol. Reps.* 5, 570.

GREEN, MARTIN (1961). *A Mirror for Anglo Saxons*. Longmans.

HEBB, D. O. (1945). *The Organisation of Behaviour*. New York: Wiley.

HOGGART, RICHARD (1957). *The Uses of Literacy*. Chatto and Windus.

JONES, W. R. & STEWART, W. A. C. (1951). Bilingualism and verbal intelligence. *Brit. J. Psychol. Statistical*, 4, 3-8.

KITCHEN, P. I. (1944). *From Learning to Earning*. Faber.

MINISTRY OF EDUCATION REPORT (1960). *Education in 1959*. H.M.S.O.

NISBET, JOHN (1953). Family environment and intelligence. *Eugen. Rev.* 45, 31-42.

PIAGET, J. (1950). *The Psychology of Intelligence*. Routledge and Kegan Paul.

THOMSON, SIR GODFREY (1945). Distribution of intelligence among University and College students. *Brit. J. educ. Psychol.* 15, 76-79.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1949). *The Trend of Scottish Intelligence*. Univ. of London Press.

Universities Quarterly (1950). Editorial comment (May).

Universities Quarterly (1952). Editorial comment (February).

VENABLES, ETHEL C. (1960). Placement problems among engineering apprentices in part-time technical college courses. Part I. *Brit. J. educ. Psychol.* 30, 237-243.

VENABLES, ETHEL C. (1961a). Placement problems among engineering apprentices in part-time technical college courses. Part II. *Brit. J. educ. Psychol.* 31, 56-68.

VENABLES, ETHEL C. (1961b). Changes in intelligence test scores of engineering apprentices between the first and third years of attendance at college. *Brit. J. educ. Psychol.* 31, 257-264.

VERNON, P. E. (1960). *Intelligence and Attainment Tests*. Univ. of London Press.

WARBURTON, F. & VENABLES, E. C. (1956). Relationship between the intelligence of Technical College students and size of family. *Eugen. Rev.* 47, 245.

Book Reviews

Pattern and Growth in Personality. By Gordon W. Allport. New York: Holt, Rinehart and Winston, 1961. Pp. xiii + 593.

This book, the author tells us, is in one sense a revision of his earlier work—*Personality: A Psychological Interpretation* (1937)—in another sense it is entirely new. The original book was as near as could be an exhaustive treatment of an inexhaustible subject, and since then the dimensions of inexhaustibility have increased, but the author remains undaunted. He has taken note of advances in learning theory, of client centered therapy, of 'ego Psychology', person perception, existentialism and phenomenology—to mention only some of the new foci of interest since 1937. It stands to reason then that this book contains discussions of many unpalatable views, but however unpalatable a theory of personality may be Allport reviews it in a kindly way and gives it a good run for its money. Except perhaps in the case of Eysenck: 'I object strongly,' he says, 'to a point of view that is current in psychology. Eysenck states it as follows: "To the scientist, the unique individual is simply the point of intersection of a number of quantitative variables". The major objection raised is that variables do not just intersect: they also interact.' Eysenck's 'simply' was perhaps an over-simplification, but it is evident that we have here a quite irreconcilable difference of opinion. Eysenck might well agree that variables do sometimes interact, but when they do so new variables emerge which also intersect in the self-same unique individual.

Allport, though well known for very many important contributions to psychology, is probably best known for his whole-hearted acceptance and elaboration of W. Windelband's distinction between *nomothetic* and *idiographic* descriptions and for his doctrine of *functional autonomy*.

Science is a nomothetic discipline in that it sets out to establish laws or generalizations. History and other 'cultural' disciplines are concerned with unique events and unique persons. Allport seems to find a problem for science in the facts of uniqueness or individuality, and a special problem for psychology in the study of personality.

There are several issues to be disentangled. The concept of 'individuality' must be distinguished from the concept of personality. The statement 'No two men possessing individuality are exactly alike in all respects', is a tautology. To say that no two persons are exactly alike in all respects is, if true, to state an important empirical fact. There is no *self-contradiction* in suggesting that two monozygotic twins reared in the same environment would be exactly alike in personality.

The empirical fact that all individuals (i.e. numerically distinct entities) have 'individuality' (i.e. differ in some respects from others of their class) is not a problem (if it is a problem at all) only for psychology. All the biological sciences have to take account of uniqueness and individuality. Almost certainly no two oak trees are exactly alike and the botanist would no doubt agree that the unique individual oak tree is a point of intersection of a number of quantitative variables (which may interact). It is odd that in the 30 years since Allport's earlier book was published (and nearly 60 years since Windelband's paper) so little progress has been made in the clarification of the issue regarding nomothetic and idiographic studies. Allport is surely right in his faith in a sensible compromise between and a rational synthesis of nomothetic and idiographic points of view. A part of the solution is found in the recognition that the biological sciences are not concerned with 'laws' of the form: If p then (invariably) q . They are more interested in statements of the form: p is a factor in the production of q . Psychologists do not expect to be able to say: All delinquents come from broken homes, or that all children who experience maternal deprivation become delinquents. What they are concerned to establish are statements about 'tendencies' or probabilities. Statements of the form $G = fx$ (where G is a biological variable, a trend or type of behaviour) assess that x is one of the factors which sometimes contribute to the production of G . And this applies to cases in which G is unique, and even when G is a changing pattern of behaviour.

The principle of 'functional autonomy' is perhaps the best known of Allport's original contributions to the theory of motivation. Chapter 10 of this book, on *The Transformation of Motives*, is a fresh and lively restatement and expansion of the original doctrine. As Gardner Murphy has well brought out in the chapter on *How we come to want what we want* in his *Human Potentialities* (1958), the concept of functional autonomy is one of a large number of overlapping concepts which include the Freudian 'cathexis', the Shand and MacDougall account of sentiment formation, the ethologists' concept of 'imprinting' and Murphy's own theory of 'canalization'. It would be a major undertaking to enquire how far these alternative formulations express differences in factual content and how far they differ simply as fancy ways of describing the same facts. It would be a major undertaking to distinguish the explanatory from the merely descriptive function of these formulations. At its simplest, the principle of functional autonomy states that activities originally engaged in as a means to an end continue to be engaged in for their own sake. This is a well established fact, an important fact but a very odd fact, but so far no one has given any clear and convincing causal explanation of it. Ultimately no doubt, the responsibility for giving such an explanation lies with the physiological psychologists, who have also the responsibility of giving a clear and convincing causal explanation of the fact of 'conditioning'—another odd fact and another concept which overlaps with the concepts of 'functional autonomy' and canalization.

Buried in this immensely erudite study of *Pattern and Growth in Personality* there are innumerable clues relevant to a doctoral thesis on the explanation of functionally autonomous activities. For a student writing a doctoral (or post-doctoral thesis) on this subject this volume is an essential text.

For whom else is it essential?

'Unlike its predecessor, the present book is intended for college students who have little or no background in psychology.' While all other changes in this new presentation of Allportian psychology can be warmly welcome it is possible to have doubts about this particular change. Of course, no student could fail to profit from contact with the mind of Gordon Allport, and with his own vivid and individual personality. But here the doubt arises. Is it a good thing to confront students with little or no background with so many of the most difficult and controversial concepts in psychology? Is it fair to expect them to make rational appraisals of the merits of the many and various schools of thought in psychology, especially when the doctrines discussed are presented by one who has a point of view of his own and writes with seductive charm? Do not these students with 'little or no background' need first of all a simple, straightforward and accurate text setting out the theories and relevant facts, written by the model text-book writer with no marked individuality, little personality and few ideas of his own?

Pattern and Growth in Personality could hardly be listed as Prescribed Reading for first year students of psychology. For one thing it is much too long. It is not so much a book as an encyclopedia. First year students could not afford to buy it, and they have not the time to read it. It is a book for more advanced students with more than 'little or no background', for students with sufficient background to appreciate a panoramic view. It is perhaps above all a book for retired professors of psychology who wish to keep in touch with current thought and more especially with the current thought of Gordon Allport. For such students this is a quite essential text.

C. A. MACE

Aspects of Psychiatric Research. D. Richter, J. M. Tanner, Lord Taylor and O. L. Zangwill (eds.). Oxford University Press, 1962. Pp. vi + 445. 63s.

Each chapter reviews progress during the last decade in an area of research. Paul Polani writes of chromosome abnormalities, H. Harris of biochemical errors, Marthe Vogt of pharmacologically active brain constituents, and R. A. Hinde on the analysis of animal behaviour. These are excellent chapters which can confidently be recommended to students of medicine or psychology.

The quality of the other sixteen chapters, and their usefulness to students, is varied. A version of John Bowlby's Adolf Meyer lecture on childhood bereavement makes one chapter.

D. A. Pond has some fresh things to say about the psycho-physiological problems of epilepsy, which will appeal to the discerning, but may be missed by the ordinary reader. J. A. V. Bates writes on the history of the E.E.G. and its place in research on cerebral function. Alick Elithorn finds grounds for distinguishing endogenous from exogenous depression, and places emphasis on the value of imipramine and leucotomy in certain types of cases. John Smythies writes about biochemical factors in schizophrenia with an optimism undimmed by the disappointment of the hopes held out in his previous articles on this topic.

Perhaps the most remarkable thing about the book is that it is addressed to medical students in their last two years, presumably as a guide to research in psychiatry. Yet none of the four distinguished editors is a psychiatrist or teaches medical students. Its title is misleading, for it is not so much about psychiatry as about areas of research from which may emerge results of relevance to psychiatry. The choice shows a bias towards areas relevant to constitutional and genetical theories, and there is much less of interest for those who find psycho-social theories of more help in the investigation and treatment of patients. The editors are members of the Research Committee of the Mental Health Research Fund, who are said to have prepared the volume. The uneven coverage of the book, and the bias towards peripheral areas, raises doubts about the Fund's appreciation of the problems of psychiatry.

D. RUSSELL DAVIS

Social Psychology Through Experiment. George Humphrey and Michael Argyle (eds.). London: Methuen, 1962. Pp. 208. 21s.

In the words of the editors 'the aim of this book is to provide a handbook and guide to practical class work in social psychology'. Under each of seven major headings three experiments are fully described; most of them can be repeated in a two-hour session, provided the instructor has devoted at least the same amount of time to the preparation of research instruments and apparatus. The topics selected and the authors of the chapters are: 'Social Perception', Henri Tajfel; 'Motivation and Conflict', Peter Robinson and Michael Argyle; 'Communication', A. N. Oppenheim; 'Small Social Groups', Thelma Veness; 'Interviews', D. R. Price-Williams; 'Social Class', Gustav Jahoda; and 'Small-scale Social Surveys', J. F. Morris.

As one would expect from this list of well-known and competent authors, the experiments are throughout described with meticulous care and clarity. The introductory comments to each chapter refer briefly and, by and large, adequately, given the limitation of space, to the general state of knowledge in the field, but the bulk of the presentation is rightly reserved for the description of methods. The level of statistical sophistication and knowledge of psychological concepts required to appreciate these experiments varies from chapter to chapter. As a result, the book must be used selectively in introductory courses, but none of the chapters should defeat an undergraduate in his second or third year. The uniformly high quality of all chapters and the authors' disciplined adherence to the main purpose of the publication is reason to congratulate the editors on their choice of contributors.

As an inevitable by-product of the plan of the book, the most difficult step in research—the formulation of a research problem for experimental work—is, of course, not dealt with. But it should not be difficult for a good teacher to use the book without giving the impression that technical competence is a substitute for hard and creative thought.

One minor criticism: While the editors, in an Introduction for Humanists and an Introduction for Psychologists, argue the case for experiments, they gloss over the ethical problems this raises too easily. It is true, the examples in the substantive chapters hardly demonstrate the ethical problems in psychological experiments. But these problems exist, even in the famous Asch experiments, and much more so in Professor Skinner's work, for example. One cannot help feeling, even as an admirer of such research, that students must be induced to take the ethics of their procedures seriously, while there is time.

Altogether, however, this book makes one wish that the number of students and instructors in social psychology will increase rapidly in the near future to make it the success it deserves to be.

MARIE JAHODA

Fundamentals of Psychology. By F. A. Geldard. Wiley: London and New York, 1962. 57s.

The author of a new textbook of psychology faces a challenge: the last few years have brought many novel contributions to be assessed and related to the more historic sources of the subject. Professor Geldard had the advantage of a solid basis for his book: he was invited to revise the last of the well known series of textbooks edited by Boring, Langfeld and Weld; nevertheless he is not altogether successful in meeting this challenge.

A series of chapters deal with motivation, learning and perception, in that order, as in the earlier book. This sequence is then repeated at a more advanced level, and the volume ends with chapters on personality, social psychology and related topics. For the most part the subjects dealt with in 1948 have been brought up-to-date by the inclusion of relevant recent work chosen in a competent and painstaking fashion, and many new illustrations have been added to illustrate the history of the subject and to amuse and interest the reader. This book is almost 200 pages shorter than its predecessor, and to some extent the selection and abridgement that has allowed this must be regretted; not infrequently subjects are more clearly treated in the older volume. There is now no simple general account of the nervous system (though the reticular system, autonomic connections, and the anatomy of the sense organs are severally described), nor of statistical terms or psychophysical methods, and even the short description of sound waves and the physics of cochlear function which previously introduced audition has disappeared.

The views and the illustrative experiments selected necessarily reflect the author's preferences; their exposition, however, suffers from a flat and platitudinous style that more often achieves vagueness than simplicity. There is little critical consideration of the facts presented and little attempt to compare and contrast different findings. Perhaps for this reason, the series of topics in some chapters appear poorly related. The latter half of 'Problem Solving and Thinking', for example, presents in succession Duncker's experiments, the 'water-jar' problem, delayed responses, images, synesthesia, hallucinations, dreams and concept formation. In keeping with the tradition of these books Professor Geldard values 'facts' highly and is suspicious of 'theory'. A consequence is that the theory presented tends to be conventional and dogmatically stated. It is for the most part very similar to that contained in the previous volume, and divergent views and new developments receive little recognition.

Psychology is defined as 'the science of human nature'. The enquiring student will discover, *inter alia*, that 'There is abundant evidence that [instincts] occur in many sub-human forms' but '[they] are missing in the analysis of human behavior'. Emotions can be analysed introspectively; pleasure and displeasure 'prove to have much in common and to consist of somewhat vaguely localized patterns of pressure sensations'. Reinforcement is related to 'the importance of pleasure and pain in shaping human conduct', and the student will find that 'a thought is but a collection of words or other symbols strung together in associative connection'. Some views are presented which seem unlikely to be widely shared. Paranoia is described both as an 'imbalance of personality' and a 'secondary motive'. The redistribution of visceral blood evoked by emotion is partly attributable to 'a squeezing action of the liver'. Dilatation of the pupil following a sudden sound 'is part of its autonomic function of signaling the widespread disturbance of "startle"'. The personality is protected by forgetting: 'If we remembered everything that happens to us ... we should probably rapidly become a confused mass of disorganized traits.'

Some readers may regret the exclusions, especially of the modern developments which constitute the 'new frontiers' of psychology. There are sections on the reticular system, intracranial self-stimulation, imprinting and teaching machines, but a subject such as skill is dealt with mainly by reference to maze and verbal learning, with no mention of information theory, of 'feedback' and the study of self-regulating devices, nor even of 'knowledge of results' or reaction times, and the only psychologists mentioned in the explanation of 'The Learning Process' are Pavlov, Thorndike, Freud and Skinner. The student who reads the sections on learning, instinct, attention, child development, thinking or hearing will find little or nothing of the views of men such as Hebb, Tolman, Tinbergen, Broadbent, Piaget, Bruner or von Békésy, while on topics such as the functions of the cerebral cortex, signal

detection theory, immediate memory, the effects of motivation on perception, or even visual acuity, he will find nothing whatever.

MICHAEL TREISMAN

Documents of Gestalt Psychology. Mary Henle (ed.). California Univ. Press, 1961. Pp. viii + 352. 64s.

This work does not, as the title may suggest, offer the reader a selection of writings representative of the entire history of a distinctive school of thought or from those works which many would regard as basic in the development in the Gestalt view of phenomena. The historically relevant papers of Cornelius, Dilthey, Husserl, Meinong, von Ehrenfels and Rubin, the important controversy between the schools of Graz and Berlin and the early writings of Wertheimer, Koffka and Köhler are, for example, not presented.

The *leitmotiv* of this readable work is conveyed by the title of the first of the twenty-two collected papers, 'Gestalt Psychology Today', a reprint of the Address of the President of the A.P.A., Wolfgang Köhler, in 1959. The papers are arranged in five groups; Essays by Max Wertheimer, General Theory, Cognitive Processes, Social Psychology and Motivation, and The Psychology of Expression and Art. In all five parts, the implications of a Gestalt approach to the wide variety of problems treated do emerge, though possibly not with optimum clarity for a younger generation of readers who might have been helped by some reference to earlier writings dealing with the fundamental principles of the school. Important, well-known empirical studies in perception by Köhler, Wallach and Asch reappear and some interesting theoretical papers, such as 'On Truth', 'Some Problems on the Theory of Ethics', 'On the Concept of Democracy', by Max Wertheimer, 'Psychological Remarks on Some Questions of Anthropology', 'Psychology and Evolution', by Köhler, and 'Emotion and Feeling in Psychology and Art' by Rudolf Arnheim, suggest further and valuable applications of the Gestalt point of view.

The present collection of papers certainly presents some of the influence of the Gestalt school upon the subsequent history of psychology; but the reader interested in tracing affinities between the basic tenets of the earlier writings of this school and more recent work, might wonder why studies by A. Ames, Jr., Attneave, Bender, Egon Brunswik, Ivo Kohler, Werner and Wapner are not more fully represented.

F. V. SMITH

Cognition. By F. H. George. London: Methuen, 1962. Pp. 309. 32s. 6d.

What Dr George is trying to do in this survey of cognition is just what we should like to see done; to synthesize relevant data from all the contributory disciplines—experimental psychology, learning theory, cybernetics, ethology, linguistics, logic and philosophy—to demonstrate how these disciplines mesh in together, and what sort of viewpoint, acceptable to them all, seems to be emerging. The prospect is an exciting one, particularly if Dr George is right in thinking that we are now 'within hitting distance of solutions' for the problems of cognition, and the molar aspects of behaviour in general.

The result, alas! is strangely unsatisfactory. 'Cognition' is an uneven and somewhat indigestible book, which fails to clarify the problems it deals with or to provide a convincing solution. Many of its hasty summaries will be too slight for psychologically sophisticated readers; its more technical accounts too compressed and difficult. It is hard to know to what class of reader Dr George is trying to communicate, and his alternations of difficulty-level may well baffle those not as widely erudite as Dr George himself.

There are, however, more serious faults. In spite of his claim to precision Dr George is at times far from precise. One wonders sometimes whether these are tape-recorded lectures. 'It was found that if the dogs were kept in a soundproof room, and a bell, gong, metronome, or other sound was made to occur at the same time as feeding took place . . .' (p. 37). At the same time? 'The distinction between cognition and conation is primarily due to Kant' (p. 24). Plato? Aristotle? St Thomas? In spite of his claim to straightforwardness Dr George is at times far from clear; he slurs over difficulties and leaves his reader

confused. 'Ideas and feelings can be the direct and immediate activity of the nervous system' (p. 200). How? This is just what we want to know, but are certainly not entitled to assume. Not dependence, parallelism, double aspects, but apparently identity is what Dr George is claiming. He should explain how this can be justified.

Then the selection of material raises queries and doubts. Certainly Dr George is entitled to select, and he has disclaimed any intention to completeness. *Cognition* is not a text-book. Nevertheless a theory which professes to be within hitting distance of a final solution must not reveal too large gaps, must leave out nothing of major significance. But there are large gaps in Dr George's account—nothing on symbolic thinking, nothing on creativity, nothing on mathematical thinking (in some ways the archetype of thought), no consideration of the 'Traumdeutung', no reference to post-1930 Piaget. Are not these matters as relevant to a theory of cognition as the learning of rats, distorted rooms, the programming of computers, and the other things Dr George considers? Perhaps the truth is that the final solution of the problem of cognition is a little further round the corner than Dr George's optimism makes out. He has given us some of the ingredients—but they hardly as yet make a cake.

L. S. HEARNSHAW

School Leavers—Their Aspirations and Expectations. By Thelma Veness. London: Methuen, 1962. Pp. 252. 25s.

'Comment is free, but facts are expensive.' You may have noticed that *Private Eye*, in a recent parody of the style and content of the *Guardian*, included this revision of C. P. Scott's famous dictum. Whatever the merits of the parody, the revision stands in its own right, and explains the current spate of fact-free comment on the state of our society. In contrast, though parts of it are unavoidably hard going, *School Leavers* is refreshingly factual. It is a team project, initiated and mainly carried out by members of the Department of Psychology of Birkbeck College, University of London. Miss Thelma Veness, who played a leading part in the project, co-ordinated and wrote up the several research reports.

The facts on which the book centres were provided, in 1956, by 1302 boys and girls, most of them in their school-leaving years. They were at schools (16 in all) in two English counties; one in the Home Counties, from which about three-quarters of them came, and one in the West of England. Three types of school were taken: secondary modern, technical and grammar. Despite welcome help from the Nuffield Foundation, the team was small, many of them part-time, and resources were extremely limited. Thus it proved impossible to arrange for an effective sampling procedure, and the relevance of the findings is therefore limited. Another unfortunate limitation was the unwillingness of educational authorities to allow the research team to study the occupational and social background of the school-leavers' parents.

What remains is, however, substantial. The school leavers appeared to be much in favour of ambition (between eight and nine out of ten of them thought it a good thing, on the whole, to be very ambitious). Their own ambitions were rather firmly cast within the existing social framework, which they saw with very little distortion from fantasy. As might be expected, their own families and schools powerfully influenced them. Very interestingly, one of the main research instruments—an autobiographical essay in which the school leavers were asked to imagine themselves, at the end of their lives, writing on what had happened to them—revealed quite a strong tendency among the boys to see their own children as following in their footsteps (nearly a third of those boys who mentioned children in these essays gave their sons the same career as themselves). The girls largely couched their ambitions in terms of husband and children, though a majority of them took it for granted that they would take jobs at some point in their married lives. Ambitions, it seems, are largely influenced by *direct* social contacts.

An attempt was made to use Riesman's typology for the classification of reasons for choosing jobs. The largest category among both boys and girls was inner-direction, with other-direction second and tradition-direction a poor third. The technical school boys and girls appeared to be the most other-directed. They also seemed to be the most ambitious, in terms of a rather complicated 'ambitiousness' scale. This scale, by the way, for which Mr Dennis Lamberth was mainly responsible, is an interesting and serious attempt to develop a technique that effectively discriminates between more and less ambitious pupils.

Professor Mace, in a characteristically sprightly and informative foreword (he was project director) praises the school leavers for their realism, and recommends the book above all to parents who have children. Miss Veness has done her very effective best to organize and present the book so that they will be able to steer easily through the detailed and complex material. The publishers have helped by a good lay-out and some pictures. But the book is also a most scrupulous and well-documented research report.

J. F. MORRIS

Measuring Human Motivation. R. C. Birney and R. C. Teevan (eds.). New York: van Nostrand, 1962.

The series of publications to which this book belongs is intended to introduce students under supervision to original writings, but the majority of such students should have access to most of the journals from which the nine articles in this volume are drawn. The articles are not supplemented by editorial comment and the unsupervised reader may well find it very difficult to integrate and evaluate the contents. Editorial reference to later theoretical and empirical work on the methods described would have been useful and brief passages to link the articles would have helped still further. As it is, this book can be really useful only to a limited section of the student body of this country.

While the articles presented are all relevant to the measurement of human motivation, they do not constitute a representative sample of the field. As the title suggests the emphasis is upon methods of measurement rather than ways of thinking about motivation, while the projective methods command most attention. Further, the projective methods described are mainly of the intermediate type—imaginative stories told to fairly structured pictures. The four articles about this method are most useful: they show the manner in which the study of individual differences among the normal population has grown out of techniques originally designed to facilitate a clinical diagnosis, with a corresponding increase of concern about objectivity, reliability, and empirical validity in an experimental context. Psycho-analysis is directly represented by Freud, Jung and Ferenczi, whose article is devoted to Freud's theory of dreams. Freud's advice on the conduct of analytic interviews is a surprising selection, but not quite so surprising as the penultimate article by Taylor in which this author tests predictions from Hullian learning theory using a questionnaire for measuring motivation. The editors seem to have felt guilty suddenly about their limited frame of reference, but the isolated inclusion of an article using a more direct method of measurement is either inadequate or inappropriate. The final chapter mentions some of the limitations of projective techniques. The earlier chapters are well supported with sample protocols and their analysis.

Thus no dominant theme is introduced, developed, and discussed. The editors set themselves an extraordinarily difficult task, and while their selections represent some excellent individual articles, the continuity can only be established by a reader already conversant with the problems of measuring human motivation.

W. R. ROBINSON

Family Environment and Delinquency. By Sheldon and Eleonor Glueck. London: Routledge and Kegan Paul, 1962.

The data for this complex work derives from its predecessors, especially the Glueck's classic *Unravelling Juvenile Delinquency*. Here, they grapple courageously with the complex nature-nurture problem of the aetiology of delinquency. Their working methods and concepts of traits and factors and their inter-relationship differ somewhat from that of most psychologists who are usually employing correlation and factor-analytical techniques when dealing with specific personality-theory models in similar circumstances (however, the authors' use of the χ^2 and interesting 'multiple comparison' techniques, later on in their book, strikes a more familiar note). The first three of five chapters deal with the Gluecks' conception of the origin of traits and they focus attention on those forty-three traits which had not been found in their earlier *Physique and Delinquency* to vary significantly in incidence

among the inborn body-types and which they think, therefore, may possibly be closer to the environmental than to the biological end of a posited biosocial continuum.

In the following two chapters, the focus is on the twenty-three traits which had been found in *Physique and Delinquency* to vary in incidence among the four body-types, and therefore had been established as closer to the biological than to the socio-cultural zone of the continuum; the question at issue being whether some of these latter may on further examination be found to be the product of a relatively equal admixture of biological and environmental influences.

The Gluecks hold that traits are made up of genetic and external influence depending on the degree of association between them, and they distinguish between social or cultural factors and physiological traits at various points in relation to opposite ends of a continuum. They advance several hypotheses about pre-existent traits which may serve as a stimulus to specific environmental responses, and about the possible association between factors and traits which may be dynamically reciprocal in their influence, namely, given the trait, the factor is likely to follow, and given the factor (as a response to the trait), the trait is likely to be strengthened.

In their finding, they differ, for instance, from Eysenck in that they place such variables as introversion-extroversion and neuroticism on the social-environmental end of the continuum instead of at the constitutional end.

The appendix, apart from the tables, contains the case history of 'Henry W.', which they had hoped to present already in one of their earlier volumes, giving it as a prototype of their now familiar approach of combining quantitative and qualitative material. What readers must decide for themselves is whether they wish to look primarily for statistical neatness and a high degree of measurable validity or whether they want to look primarily for insightful guidance with reference to many clinical entities in the pioneering works of the Gluecks.

R. G. ANDRY

Industrial Psychology. By J. Munro Fraser. Oxford: Pergamon Press, 1962. Pp. ix + 181. 12s. 6d.

Is industrial psychology essentially a technology which uses psychological techniques, such as rating scales, mental tests, etc.? Or is industrial psychology a field of enquiry, comparable with social psychology or experimental psychology? The implication of this book is that industrial psychology is a bit of both. The author, having introduced the readers to the historical roots of the subject, promptly gets down to considering the theoretical foundations of modern industrial psychology. In particular, he sets out to examine human needs and their satisfaction. He adopts a classification of needs which is relatively non-controversial; he then considers work as a means of need satisfaction. This leads to a discussion of 'frustration-instigated' behaviour in the industrial situation. This approach to human problems in industry is presented with clarity in a non-technical manner. Perhaps some indication of other possible approaches would have helped to balance the picture.

The social psychology of industry receives in the book a very adequate treatment. Two chapters out of ten are devoted to the many problems of the matching up of people with jobs; the methods of job specification and the assessment of the individual are well explained in practical terms. Other preoccupations of industrial psychology, such as training, work study, fatigue, are considered more briefly. Little mention is made of the numerous experimental studies that underlie the contribution that psychology makes to the design of industrial equipment. Having started with an outline of problems and aims, the book ends with an attempt to evaluate the contribution of industrial psychology to knowledge and human welfare.

References are found only in footnotes. A list or lists of books and articles suggested for further reading would have been helpful to some readers; but this is perhaps not a serious omission. There has been for some years a real need for a modern introduction to industrial psychology; this useful book has met the need satisfactorily.

W. SLUCKIN

Genetic and Environmental Factors in the Development of 'Primary Reaction Patterns'*

By MICHAEL RUTTER,† SAM KORN‡ AND HERBERT G. BIRCH§

In the course of a longitudinal study of 128 children, individual styles of behaviour have been delineated and followed through successive age-periods. The contribution of genetic forces to the formation of such reactive patterns has been examined by the study of 56 sibs and twins who form part of the study population. Various non-genetic influences were examined through a longitudinal analysis of stability and instability over time of the characteristics of reactivity for the whole sample of sibs and twins. The relevance of the findings for a theory of individuality in psychological development is considered.

In the course of longitudinal study of 128 children it was found (Chess *et al.*, 1959; Thomas *et al.*, 1960; Birch *et al.*, 1962) that individual behavioural styles could be delineated in the first months of life. These styles, defined in terms of activity level, threshold of responsiveness, rhythmicity of functioning, adaptability, intensity, approach-withdrawal, mood, persistence, and distractability not only distinguished one infant from the other, but persisted as characteristics of individuality in subsequent years. Although the individual styles were termed 'primary reaction patterns' because of the early age at which they could be delineated, the adjective 'primary' in no sense implied innate or simple genetic determinance of the observed individual differences. Rather, on philosophic grounds (Birch, 1961; Lehrman, 1953; Schneirla, 1957, 1959; Thorpe, 1960), an interactionist position was preferred. The behaviour patterns were therefore viewed as the result of the interaction of organismic and environmental factors.

The existence of both twins and sibships in the population of children studied made it possible to explore the relative contributions of genetic and non-genetic factors in the origin and persistence of individuality of behavioural style. Secondly, since there is some instability in individual behavioural patterning over time it was possible, through the analysis of constancy and variability in behavioural pattern, to seek out developmental factors which contributed to the patterning of behaviour in successive age-periods. These complementary explorations can add to our understanding of genetic and environmental factors in the development of behavioural individuality.

Twins and sibs

The comparison of like-sexed monozygotic and dizygotic twins has been regarded as a natural experiment allowing the differentiation of genetic and

* An analysis proceeding as part of a longitudinal study of child development, supported in part by the National Institute of Mental Health, Grant No. M-3614.

† Nuffield Medical Travelling Fellow, now at Medical Research Council, Social Psychiatry Research Unit, Institute of Psychiatry, Maudsley Hospital, London.

‡ Department of Psychology, Hunter College, New York City.

§ Department of Pediatrics, Albert Einstein College of Medicine, New York City.

environmental influences (Slater & Shields, 1953). It has been argued that any differences found between monozygotic twins must be due to non-genetically determined effects. The view has also been advanced that one may act on the additional assumption that the environments of both varieties of twins will show similar intra-pair differences. Thus it has been reasoned that the difference in concordance between monozygotic twins, on the one hand, and dizygotic pairs, on the other, allows an estimate of the degree to which genetic influences determine the variables studied.

It is necessary to examine the assumptions involved in this argument. The statement that any difference found within pairs of monozygotic twins must represent the effects of non-genetic factors is unarguable for most practical purposes. On the other hand, the second assumption is less straightforward. First, non-genetic effects may still be constitutional in so far as influences during foetal life are important. Not only is there some evidence that prenatal and perinatal factors may have a considerable influence on behaviour (Pasamanick & Knobloch, 1961), but these factors usually act to produce greater differences within monozygotic pairs than within fraternal pairs. For example, birth weight differences are greater in monozygotic twins and foetal mortality is higher (Newman *et al.*, 1937). These facts suggest *greater* intrauterine environmental discrepancies within the monozygotic pairs. A consideration of post-natal course also fails to support the second assumption. It has been argued that the range of post-natal environmental differences is similar within both types of twin-pairs (Slater, 1952). However, the contrary point of view has been well put by Jackson (1960), who stressed the unusually close and distinctive relationship between monozygotic twins (especially females) and its possible implications for development. Fraternal twins are much less likely to be dressed alike, a fact which may reflect a less uniform parental attitude towards dizygotic pairs. On the other hand (and this has been evident in our sample), parents of 'identical' twins sometimes give great emphasis to small differences.

In addition to the difficulties inherent in the logic of twin analysis, two problems of data treatment and interpretation remain to be considered. In previous studies of twins, examinations of concordance within pairs has been made with full knowledge of which pairs were monozygotic. This allowed the influence of biases which can greatly distort the results (Jackson, 1960). It has been possible in the present study to avoid such contamination to some extent. The scorers of the behavioural data were unaware, in most instances, that the child being scored was a twin. Further, all data were collected prior to the determination of the mono- or dizygotic character of the twin pairs.

Finally, in assessing the influence of environment by the study of twin pairs, the range of environmental discrepancy is limited to the differences between the environments of two like-sexed sibs nurtured within the same home. Although there is evidence that parental attitudes and practices may be markedly dissimilar for different children within the same family (Stott, 1941), it is hardly likely that such differences approach those found between different families in the general population.

Clearly, the use of twin comparisons is not without difficulties. Nevertheless, the study of twins provides the best opportunity to examine the interaction of genetic and environmental influences. Many problems disappear if one can study monozygotic twins separated from birth. However, such opportunity is rare and only one such pair is included in the present sample.

METHOD

The overall longitudinal study included eight pairs of twins and 26 pairs of sibs, all of whom at the time of the present analysis were at least 2 years old. All pairs of twins were carefully examined and detailed notes made of those physical characteristics used for twin designation (Gedda, 1961). These included colour and texture of hair, distribution of head hair and position of whorl, presence/absence of hair on limbs and its colour, form and colour of eyebrows and lashes, colour of eyes and pattern of iris, shape of features (nose, ears, mouth, etc.), skin colour and texture, teeth (size, shape, whether even or uneven), body build and general appearance. For a pair to be regarded as monozygotic, both twins had to be closely similar on all these characteristics. Height was measured and pairs of twins differing by 1 inch or more were regarded as dizygotic. Weight was also recorded but not used to determine zygosity. Fingerprints were taken and a ridge count made according to the method described by Holt (1961). The total ridge count intra-pair difference of the monozygotic twins did not exceed seven in any instance. The range of ridge-count discrepancies within dizygotic pairs was 18-24. In view of the type and age of the sample studied and the likelihood that permission would be withheld in some instances, blood samples were not taken for determination of blood groups. However, although this would have provided a more definitive differentiation, in no case was there any reasonable doubt as to the type of twin-pair. In all, three monozygotic pairs, five dizygotic pairs, and 26 pairs of sibs were studied.

The prime sources of data utilized in the present analysis have been detailed narrative accounts by the parents of the child's behaviour over time in his everyday life. Such accounts were obtained several times yearly. The rationale of the approach, the method of scoring protocols, the categories of behavioural styles examined and measures of reliability and validity have all been dealt with in previous reports (Chess *et al.*, 1959; Thomas *et al.*, 1960; Chess *et al.*, 1960). Scoring for any category was according to a 3-point scale. The nine characteristics rated were *regularity*, *adaptability*, *mood*, *intensity*, *approach/withdrawal*, *persistence*, *distractability*, *activity*, and *threshold*. For such a category as *regularity* the three points would be regular, variable (i.e. sometimes regular and sometimes irregular), and irregular rhythm of functioning. The scores for all categories were so distributed that there was a pronounced skew towards one of the polar scoring positions and, in all but the last two categories, the middle position was rarely used. Accordingly, each category had a most frequently occurring or *modal* score position, and a least frequent or *amodal* score position.

In order most effectively to consider individual differences in behavioural functioning, a ranking model was employed to increase the range of variation in the sample. By this method the children were characterized in terms of the relative degree to which they exhibited a particular aspect of any of the categories of behaviour. The child's score represents the ratio of the number of occasions on which his behaviour has been rated to fall at one of these scoring positions to the total number of his behaviours that were scored for this category. Thus, if in a given child with a total of 20 behaviours in the *regularity* category, four of these were rated as irregular, then the per cent score for irregularity would be $4/20 \times 100$ (i.e. 20 per cent). Obviously such a score could be computed for any score position for any behavioural category. In order to treat categories uniformly the position chosen for the categories of *adaptability*, *regularity*, *intensity*, *mood* and *approach/withdrawal* was the polar score position which occurred least commonly in the group of children as a whole—that is the *amodal* score position (unadaptability, irregularity, high intensity, negative mood and withdrawal from new situations respectively). The children were next ranked in order of the magnitude of these per cent scores. This procedure results in a series of ranks from 1 to 92, the child's rank representing his position in this ranking sequence.

In the categories of *activity* and *threshold*, unlike the others, all three scale positions were

frequently used and it was thought desirable to use a technique which took into account the distribution of behaviours across all three points. For this reason a weighted score method was utilized. The positions low, medium and high were given weights of 1, 2 and 3 respectively, producing a score having a minimum of 1 and a maximum of 3. Thus in any one child, if 2, 3 and 5 items of behaviour occurred in the scale positions of low, medium, and high activity, these would be multiplied by 1, 2 and 3 respectively, to give derived scores of 2, 6 and 15. The sum of derived scores (23) is then divided by the original sum of scores (10) to produce a weighted score of 2.3. Children were then ranked according to the magnitude of their weighted score.

Although all three scale points were often scored in the categories of activity and threshold, in the other categories one scale point was scored less often than the others. Much of the variation among the children occurred in the frequency with which this least common, or amodal, scale point was used. This variation was to some extent obscured by use of the weighted score. To counteract this, a modification was employed wherein the weighted scores were derived not from the child's raw score but rather from the raw score adjusted to give equal weight to all three points. Thus, if, in the group as a whole, low, medium and high activity scale position scores occurred in the ratio of 1 : 3 : 4, then the scores for each *individual* child would be multiplied by 8/1, 8/3, and 8/4 respectively to produce adjusted scores. These adjusted scores were then treated to produce a weighted score by which the children were ranked.* Low frequency of scored items in the categories of *persistence* and *distractability* did not permit the analyses described and for present purposes these categories were omitted.

RESULTS

If any characteristic has a largely genetic basis, then, other things being equal, three consequences follow: (1) monozygotic pairs will be more alike than dizygotic, (2) dizygotic pairs will be no more alike than sibs, (3) no large differences should be found within monozygotic pairs. The mean rank differences within pairs of monozygotic twins, pairs of dizygotic twins, and pairs of sibs (at the same ages) are shown in Table 1. It can be seen that, in general, monozygotic pairs were more alike than dizygotic (17/21 instances, $p = 0.004$ sign test, one-tailed) and that there were no significant differences between dizygotic pairs and pairs of sibs. It was also calculated that pairs of sibs were usually no more alike than pairs of unrelated children. These general relations are supported by a series of more detailed comparisons.

The group of twins and sibs were examined to test the three consequences which flowed from the assumption that a characteristic had a genetic basis. The first consequence was fulfilled for *adaptability*, *activity* and *withdrawal* for all three age-periods (first, second and third years of life) and for the remaining behavioural categories in two out of three age-periods. In the first year monozygotic pairs were more alike than dizygotic pairs for all seven categories. In the second year for four categories, and in the third year for six categories. In no year were the dizygotic twins more alike than the monozygotic for any category.

In order to test the second consequence, that dizygotic pairs will be no more alike than sibs, it was seen how often differences within fraternal pairs more nearly approximated to those within sib pairs than those within monozygotic pairs. In

* This somewhat elaborate method was employed in an attempt to use a score making the maximum use of the information available. Studies made since completion of this paper have shown that the simpler per cent rank score used in the other five categories provides comparable results. Rank correlations between the two methods of scoring in the categories of *activity* and *threshold* exceed 0.80 at all age periods. For this reason, in subsequent analyses the per cent score method has been employed for all categories.

the first year this expectation held in six out of seven instances. In each of the two succeeding years it held in four out of seven instances. However, for only one category of reactivity, *approach/withdrawal*, did it hold in all three age-periods. For *adaptability*, *mood* and *activity* it held for two out of three age-periods.

Table 1. Mean rank differences within pairs of twins and sibs

Category	Year I			Year II			Year III		
	MZ	DZ	Sibs	MZ	DZ	Sibs	MZ	DZ	Sibs
Intensity	11.3	16.6	15.2	12.0	21.4	20.1	8.8	8.1	11.8
Threshold	6.8	18.5	21.3	12.2	10.4	22.2	3.0	4.2	10.0
Mood	8.2	18.1	22.3	18.0	17.9	20.2	4.2	10.1	9.7
Adaptability	12.2	14.2	16.5	13.7	22.6	18.1	3.0	8.8	11.5
Regularity	14.3	15.5	20.0	8.2	2.1	21.6	3.5	8.1	12.7
Activity	12.2	26.1	22.4	5.5	13.0	18.4	4.5	6.1	12.1
Withdrawal	8.3	18.8	13.3	15.8	23.8	19.2	6.5	11.9	13.7
No. in subgroup	3.0	5.0	26.0	3.0	5.0	26.0	2.0	4.0	16.0
No. of ranking	56.0	56.0	56.0	56.0	56.0	56.0	38.0	38.0	38.0

In testing the third consequence, that no large differences would exist within monozygotic pairs, one-third of the distribution of all scores was used as the criterion measure. It was found that only for the categories *intensity* and *threshold* did no intra-pair differences exceed one-third of the distribution of scores for the whole group in any of the three age-periods.

Table 2. Categories in which each genetic consequence is fulfilled

Consequences		
1	2	3
Withdrawal	Withdrawal	
Activity	(Activity)*	
Adaptability	(Adaptability)*	
	(Mood)*	
		Intensity
		Threshold

* Consequences fulfilled in two out of three age-periods only.

Table 2 shows those categories in which the evidence from examination of each of the three different consequences suggested a strong genetic loading. The table suggests that for no category of reactivity do all methods agree in indicating a preponderantly genetic basis. The strongest evidence for a genetic component was present for *activity*, *approach/withdrawal* and *adaptability*. *Intensity*, *threshold* and *mood* appeared to have a genetic basis in only one of the three analyses, and by no method of analysis did *regularity* appear to be genetically determined. It should

also be noted that the evidence for genetic influences was stronger in the first year of life than in either of the two subsequent years.

An opportunity to examine the possible influence of non-genetic factors was afforded by the existence of special circumstances in the life histories of four of the eight twin pairs. Two of these pairs (A & B) were monozygotic and were treated as such by their parents. The mother of pair B, however, although treating the twins very similarly, was concerned with whether or not it would be better to emphasize their differences by differential treatment. One pair (C) was monozygotic but the twins were separated at birth and brought up in different families fully isolated from one another. Although each family knew of the existence of a twin, neither had any knowledge of the placement. The last pair (D) was indubitably dizygotic, but for the first 3 years the twins had been regarded as 'identical' by their parents. They had been dressed alike and the treatment of each by the mother was so similar that if one became wet both were changed.

If parental handling influenced the patterns of behaviour studied, then one would expect that the monozygotic twins reared apart would be closely alike initially but would diverge increasingly as they grew older. The fraternal pair reared as 'identical' would be expected to start less alike but should converge. The monozygotic pairs 'A' and 'B' would be expected to start alike and to remain alike.

Table 3. *Contrast analysis of twin pairs mean rank differences within pairs*

Year	Twin pairs			
	A	B	C	D
1	9.4	11.6	10.6	16.1
2	8.9	11.2	16.4	5.5
3	1.9	8.0	—	9.1

Key: A = Monozygotic twins reared in same household.
 B = Monozygotic twins reared in same household.
 C = Monozygotic twins reared apart.
 D = Dizygotic twins reared as 'identical'.

Table 3 summarizes the results of the analysis of these inferences. As may be seen from the table, in the first year the rank differences between members of each of the three pairs of monozygotic twins are all less than the mean rank difference found between the dizygotic twins reared as identical. By the end of the second year the dizygotic twins reared as identical have markedly reduced their mean rank difference, whereas the mean rank difference on the monozygotic individuals separated at birth and reared apart has increased to a level equal to that found in the dizygotic pair at the end of the first year. The monozygotic twins reared in one household started out with low mean rank differences and continued to have only small differences throughout the first 3 years. Unfortunately this set of comparisons must remain semi-quantitative because the small number of cases precludes more detailed statistical analysis.

Stability of reactive patterning in the first three years

The analysis of the genetic factors contributing to the persistence in individuality in reactivity indicates that, although the genetic factors account for part of the pattern, a considerable contribution is being made by other influences. It is possible to identify some of these non-genetic contributions by means of a longitudinal

analysis of the stability and instability over time of the characteristics of reactivity of the sample of sibs and twins.

Table 4 shows the correlation coefficients (*rho*) between years I and II, I and III, and II and III for each of the seven categories of reactivity. It may be noted that five of the seven categories are stable between the first and second years.

Table 4. Reaction pattern stability by ranking of scores

Category	Periods			
	Correlation (rho) I-II (56 cases)	Correlation (rho) I-II (38 cases)	Correlation (rho) II-III (38 cases)	Correlation (rho) I-III (38 cases)
Intensity	0.46*	0.42*	0.12	-0.02
Threshold†	0.36*	0.36*	0.40*	0.35*
Mood	0.34*	0.44*	0.21	0.04
Adaptability	0.30*	0.35*	0.22	0.06
Regularity	0.27*	0.39*	0.38*	0.08
Activity†	0.24	0.41*	0.10	-0.07
Withdrawal	0.23	0.08	0.08	0.01

* Statistically significant at 5 per cent level or better.

† Using weighted scores.

Although correlations are positive for the other two categories, *activity* and *approach/withdrawal*, each just fails to reach the 5 per cent level of confidence. Two of the categories, *threshold* and *regularity*, are significantly stable between the second and third years, while only *threshold* is stable from year I to year III. These findings lead to exploring the sources for stability and instability over time.

Table 5. Correlations (*rho*) between categories for year I

	Intensity	Threshold	Mood	Adaptability	Regularity	Activity	Approach/ Withdrawal
Intensity							
Threshold	-0.43*						
Mood	0.48*	-0.24					
Adaptability	0.42*	-0.32*	0.36*				
Regularity	0.38*	-0.12	0.19	0.24			
Activity	-0.05	-0.08	-0.36*	-0.19	-0.14		
Approach/ Withdrawal	0.40*	-0.14	0.50*	0.68*	0.15	-0.16	

* Statistically significant at the 5 per cent level or better.

One potential source for instability over time may be that similarly named functions at two different ages may be based upon quite different behaviours, and that similar behaviours at two ages may reflect different underlying functions. One approach to this question may be pursued by examining the intercorrelations between categories at different ages. Changes in intercorrelations would accord with the view that similarly named categories are not entirely comparable at different points. As may be seen from Tables 5 and 6, the intercorrelations in the first year are often high (up to 0.68 between *adaptability* and *approach/withdrawal*). *Activity* and *threshold* are relatively independent of the other categories (apart

from a correlation of -0.43 between *intensity* and *threshold*) but *intensity*, *mood*, *adaptability*, *approach/withdrawal*, and (to a lesser extent) *regularity* show sizeable positive intercorrelations. However, by the second year the correlations are lower except for those between *intensity* and *threshold* (-0.56) and *intensity* and *adaptability* (0.52). There is some evidence that similarly named categories at different ages are not entirely comparable.

Table 6. *Correlations (rho) between categories for year II*

	Intensity	Threshold	Mood	Adaptability	Regularity	Activity	Approach/ Withdrawal
Intensity							
Threshold	-0.56^*						
Mood	0.29^*	-0.16					
Adaptability	0.52^*	-0.26	0.25				
Regularity	0.24	-0.16	0.12	0.03			
Activity	0.18	0.11	-0.28^*	0.21			
Approach/ Withdrawal	0.07	-0.20	0.28^*	0.28^*	-0.15	-0.23	

* Statistically significant at the 5 per cent level or better.

A second basis for instability may lie, not in any specific source of unreliability in any given category, but rather in the possibility that *instability* itself is a particular characteristic of individual functioning. The existence of such a possibility is suggested by the fact that within each of the nine categories there is no single curve which can encompass the distribution in changes in rank from one age-period to another. Many children show little change, while others show large swings, moving from one extreme to the other.

To test this concept the data were examined to determine whether or not the children who showed marked instability of one category were also those unstable on others. All children were ranked according to the change in rank between the first and second years, and the ranking of instability across the seven categories compared by means of Kendall's concordance coefficient (W). The coefficient obtained ($W = 0.20$) shows the relation, although weak, to be significant at the 3 per cent level of confidence, supporting the hypothesis of a general instability 'factor'. Three further questions were thereby raised. Is the postulated instability 'factor' stable over time? To what extent does it have a genetic basis? What environmental influences may affect it?

Similar analysis for the rank changes between years II and III produced a W of 0.15 which, with the smaller number of children (38) at these ages, failed to reach significance at the 5 per cent level. However, this coefficient, although less than at the first two years, was not greatly dissimilar. Therefore, children were ranked according to their general instability, using R_j as determined in the calculation of W (Siegel, 1956). Calculation of ρ between ranking of instability at the first age-period (years I and II) and second (years II and III) produced a coefficient of 0.32 which was significant at the 5 per cent level. Thus those children who showed marked changes in their patterns of behaviour at one age, to some extent, were also those showing marked changes at a later age (at any rate up to the age of 3 years).

Comparisons between pairs of twins and sibs gave no suggestion that the instability was gene-determined. That factors in the home might be relevant was suggested by the finding that the monozygotic twins, separated at birth and reared in different homes, differed to a greater extent than any other pair of twins and only one pair of sibs showed as great a difference.

Another possible source of instability derived from the fact that, for some children, information of all ages was obtained by a single interviewer, whereas information concerning other children was obtained by more than one interviewer. The possibility therefore existed that the instability of the children's behaviour might be a function of change in interviewer in successive periods. Table 7 shows the mean instability ranking of children with one interviewer only and of those with multiple interviewers for periods I-II and II-III. Differences for period I-II were negligible and at period II-III children having only one interviewer were somewhat *less* stable. Thus there was no reason to suspect that instability of scores was determined by change of interviewer.

Table 7. *Number of interviewers and stability of behaviour*

Period	One interviewer only		More than one interviewer		Total No. of Children
	Mean ranking in instability	No. of Children	Mean ranking on instability	No. of children	
I-II	27.1	(34)	30.7	(22)	(56)
II-III	24.2	(16)	16.0	(22)	(38)

Sex differences, too, might underlie instabilities. Comparison of the mean instability ranking of boys and girls, both with regard to the general factor and with regard to the individual categories, showed that sex differences at these ages made no significant contribution (Table 8).

Table 8. *Ranking of instability by sex (Periods I and II)*

Category	Mean rank on instability	
	Male (30 cases)	Female (26 cases)
Overall ranking	26.6	30.7
Regularity	30.0	26.8
Adaptability	28.8	28.1
Approach/Withdrawal	29.5	27.3
Intensity	27.9	29.2
Mood	26.9	30.3
Activity	26.9	30.3
Threshold	25.6	31.5

Specific instabilities might also follow from certain aspects of parental behaviour. For example, children in whom there is the initial characteristic of negative response to new situations might later fail to show such a pattern if the parents so structured life that a minimum of fresh demands were involved. Comparison of scores according to all situations and new situations only might demonstrate such effects and re-emphasize those individual characteristics which may have become blurred through an artificial homogenization. Preliminary examination of 20 cases using responses to novel situations only has suggested that greater stabilities of reaction patterns are found than when total behaviour is scored (Birch *et al.*, 1962).

DISCUSSION AND CONCLUSIONS

By an intensive examination of a group of 56 sibs and twins followed from birth during the course of a longitudinal study, an attempt has been made to begin to unravel some of the factors involved in the development of 'primary reaction

patterns'. It appears that, although there was some evidence for genetic influences in most categories, there were strong suggestions also that non-genetic factors played an important role. This finding is, of course, one that might apply to almost any human function. The interest lies rather in differences found between categories and on any light these may throw upon patterns of interaction between genetic and non-genetic forces.

Perhaps one of the most striking findings is that those categories (*activity* and *approach/withdrawal*) in which the evidence for a genetic basis is strongest are also those showing the greatest instability over time. Further, the two most stable categories (*threshold* and *regularity*) were those in which the evidence for a genetic influence was least strong. Although the findings regarding genetic influences must be regarded with considerable reserve in view of the small numbers of twin pairs, the findings regarding stability are reliable.

The behaviour from which scores were derived was that of the child in his natural setting and was molar in character. Therefore the finding that *regularity* of functioning appeared to be largely determined by non-genetic factors does not mean that genetic factors related to regularity do not exist. Rather it implies that (within this group), considering molar behaviour, such influences were obscured by the greater effect of 'environmental' variables. Examination of molecular behaviour might have led to different conclusions. The variation regarding *regularity* within our group was small and what small heterogeneity existed seemed to be chiefly accounted for by non-hereditary factors. If a pathological population had been chosen, different results might also have been found.

However, it should *not* be assumed that the relevant non-genetic factors were necessarily those related to the parents' child-care practices or other variables within the home. That children could be categorized according to *threshold* and *regularity* soon after birth, and that such categorization remained relatively stable, suggests that 'constitutional' differences existed between children. As, apparently, these were not genetically determined, they may have been related to pre- and paranatal influences. Such influences might be particularly apparent in a group of twins. It is well known that a number of complications of pregnancy are more common in association with multiple births. Twins are usually of less than average birth weight, and foetal mortality is higher than among singletons. Such differences are also most marked within monozygotic pairs. Whether such influences are the relevant ones and whether the differences can be regarded as normal variations related to the intrauterine environment, or whether they are better regarded as morbid characteristics along a continuum of reproductive causality (Pasamanick & Knobloch, 1961) cannot be determined from our data.

The reason for the marked instability of those categories with, apparently, the greatest genetic loading is not immediately clear. That a characteristic exhibits instability in early childhood does not mean, however, that it is lacking in developmental importance. Functions readily identifiable in infancy may change their form or cease to be recognizable in later childhood, continuity being yet demonstrable by high correlations with other derivative or related variables. Such a situation in regard to activity level was suggested by findings in the Berkeley study

where this characteristic in the infant showed considerable inconsistency by the age of 3 years, but was significantly (and negatively) correlated with later task-oriented functions (Schaefer & Bayley, 1963). A similar effect was found, too, in the Fels study, where variables of maternal behaviour had little statistical relationship with the child's behaviour at the same age, but were correlated with characteristics in later age-periods, demonstrating what they termed 'sleeper-effects' (Kagan & Moss, 1962). Correlations, by their nature, can never prove causal relationships and particular caution is necessary when one is drawing conclusions from isolated correlations, however highly significant statistically, especially when these are extracted from matrices of non-significant correlations. Nevertheless, the delayed effects that longitudinal study may show to be produced by characteristics in infancy can have considerable theoretical and practical importance.

Our failure to find significant sex difference effects when other longitudinal studies have shown that boys and girls differ markedly both in the degree to which they exhibit stability of functioning and in the relationship they show between different functions (Kagan & Moss, 1962; Schaefer & Bayley, 1963) represents only an apparent contradiction. Such differences have been either absent or less prominent in the first 3 years of life, so that the present negative findings are in agreement with other longitudinal studies.

The finding that consistency or inconsistency of functioning is to some extent a general and persisting characteristic of the child suggests that children can be meaningfully categorized according to this variable. When examining relationships between functions at different ages, it may be that these should be considered separately for those children exhibiting stability of functioning and those showing instability.

In considering the development of 'primary reaction patterns' we must emphasize that we do not regard them as things or entities. Rather, in our usage, they are terms which label complexes of functions and so serve to characterize styles of behaviour. Certain aspects of such styles have been selected for study, but clearly there are other aspects. Further, if intercorrelation patterns show that the same styles of behaviour are better described in other terms or may be more parsimoniously delineated by fewer categories, then such modifications will be adopted.

It should also be stated that 'primary reaction patterns' do not represent the intra-organismic side of the nature-nurture controversy. Any person's behavioural characteristics are derived from an interaction between his organic sub-stratum, including certain genetic potentialities, and a variety of environmental influences, some of which are essentially structural in their effects and others of which are influential through learning and other related mechanisms. Such interaction can be viewed as starting at the time of conception and continuing throughout the life-span. Actual behaviour is the outcome of the further interaction of such reactive characteristics and the effective environment at the time. 'Primary reaction patterns', then, are abstractions used to describe the characteristic styles of such interaction as shown by the individual through time.

In the broadest sense it is essential that one avoid the sterile genetic versus environmental categorization in developmental research. Clearly, effective environ-

ment, representing those features of the surround which may influence developmental course, is the result of the selective sensitivity and capacities for response of the organism. Thus, appropriate concern in development must be directed at delineation of the elements of a process of mutual and continuous interaction between organism and environment. Our attempt to explore genetic influence has occurred within this conceptual frame.

APPENDIX: Definition of Categories

1. *Activity Level*: The motor component in behaviour; the frequency with which a child shows high, moderate or low levels of motor activity in different situations. Data on motility during bathing, eating, playing, dressing and handling, as well as information concerning the sleep-wake cycle, reaching, crawling and walking, are used in scoring this category.
2. *Rhythmicity*: The predictability and/or the unpredictability in time of any function. This can be analysed in relation to the sleep-wake cycle, hunger, feeding and elimination schedule.
3. *Approach or Withdrawal*: The nature of the response to a new stimulus, be it a new food, new toy, or new person.
4. *Adaptability*: Responses to new or altered situations. This is concerned not with the nature of initial responses, but with the ease with which these are modified in desired directions.
5. *Intensity of Reaction*: The energy level of response, irrespective of its quality or direction.
6. *Threshold of Responsiveness*: The intensity of stimulus necessary to evoke a discernible response, irrespective of the specific form that the response may take, of the sensory modality affected. The behaviour involves reactions to sensory stimuli, environmental objects and social contacts.
7. *Quality of Mood*: The amount of pleasant, joyful and friendly behaviour, as contrasted with the reverse.

REFERENCES

BIRCH, H. G. (1961). *Amer. J. Orthopsychiat.* 31, 267.

BIRCH, H. G., HERTZIG, M., CHESS, S. & THOMAS, A. (1962). Annual Meeting, American Orthopsychiat. Assoc. Los Angeles, March, 1962 (Unpublished).

CHESS, S., THOMAS, A. & BIRCH, H. (1959). *Amer. J. Orthopsychiat.* 29, 791.

CHESS, S., THOMAS, A., BIRCH, H. & HERTZIG, M. (1960). *Amer. J. Psychiat.* 117, 434.

GEDDA, L. (1961). *Twins in History and Science*. Springfield, Ill.: Chas. C. Thomas.

HOLT, S. (1961a). *Brit. med. Bull.* 17, 247.

HOLT, S. (1961b). In *Recent Advances in Human Genetics* (ed. Penrose L.). pp. 101-119. Boston: Little Brown & Co.

JACKSON, D. D. (1960). In *The Etiology of Schizophrenia* (ed. D. D. Jackson). New York: Basic Books.

KAGAN, J. & MOSS, H. A. (1962). *Birth to Maturity*. New York: John Wiley.

LEHRMAN, D. S. (1953). *Quart. Rev. Biol.* 28, 337.

NEWMAN, H. H., FREEMAN, F. N. & HOLZINGER, K. J. (1937). *Twins: A Study of Heredity and Environment*. Univ. of Chicago Press.

PASAMANICK, B. & KNOBLOCH, H. (1961). In *Prevention of Mental Disorders in Children* (ed. Caplan G.). New York: Basic Books.

SCHAEFER, E. S. & BAYLEY, N. (1963). *Maternal Behavior, Child Behavior, and Their Inter-correlation from Infancy through Adolescence* (monograph in press).

SCHNEIRLA, T. C. (1957). In *The Concept of Development: an Issue in the Study of Human Behavior* (ed. Harris, D. B.). Univ. of Minnesota Press.

SCHNEIRLA, T. C. (1959). *Comparative Psychology*. Encyclop. Britannica.

SIEGEL, S. (1956). *Non-parametric Statistics*. New York: McGraw-Hill.

SLATER, E. (1953). *J. ment. Sci.* **99**, 44.

SLATER, E. T., and SHIELDS, J. (1953). *Psychotic and neurotic illnesses in twins*. M.R.C. Spec. Rep. No. 278. London: H.M.S.O.

STOTT, L. H. (1941). *Character & Pers.* **10**, 140.

TANNER, J. M. (1960). In *Human Growth* (ed. Tanner, J. M.). New York: Pergamon Press.

THOMAS, A., CHESS, S., BIRCH, H. & HERTZIG, M. (1960). *Comprehensive Psych.* **1**, 103.

THORPE, W. H. (1961). *Annu. Rev. Psychol.* **12**, 27.

Manuscript received 20 July, 1962

Early Experience and Behaviour Disorders in Severely Subnormal Children*

By MARY WOODWARD†

University College, Swansea

A previous study of the relation between early experience and subsequent social behaviour in severely subnormal children has been followed up with a similar study of 130 more advanced severely subnormal children living in hospital. Groups which were distinguished on the basis of the type of social response to the investigator differed significantly in other aspects of behaviour both in the individual situation and in the classroom. The type of social response shown by children in hospital was significantly related to the presence or absence of adverse factors in the home before admission and to the number of previous hospital admissions. The theoretical and practical implications of the results are discussed.

1. INTRODUCTION

The last two decades have seen considerable psychological study of backward adults, including those who used to be called imbeciles. This has led to the investigation of severely subnormal children, though mainly of their cognitive functions and learning abilities; a parallel interest in their social and emotional behaviour has been less in evidence. Yet the study of these aspects is equally important, both theoretically and practically: it may, for example, throw light on the origin of behaviour disorders in severely subnormal children, with practical implications for their modification.

A previous study by the writer (Woodward, 1960) described the social responses of 90 severely subnormal children living in hospital whose intellectual development was similar to that of normal infants up to the age of 2 years. Children who reacted to a social approach with distressed or avoiding reactions were significantly differentiated from the rest of the group by their pre-admission experience: 73 per cent of them, compared with 30 per cent of the others, had lived in material or emotional circumstances which might be described as adverse. They also differed significantly in their initial response to toys and in the focus of their activities.

This study has been followed up with a further one of more advanced severely subnormal children whose intellectual development is approximately that of normal children of 2-6 years. The first part of this enquiry aimed to observe the social behaviour of the children in an individual situation with the investigator, to compare this with reports of social behaviour in the classroom, and to examine age trends. The second part investigated the hypothesis, derived from the previous study, that the type of social response observed in hospital would be related to

* Based on a paper read to the English Division of Professional Psychologists, February, 1962.

† Previously at the Fountain Hospital, London.

social and emotional experiences before admission. The study was restricted to children without physical handicaps, since these might also contribute to behaviour disorders.

2. SUBJECTS

The subjects were living in a long-stay hospital for severely subnormal children in London. From all the children attending the training centres, those with a known visual or auditory defect or severe motor disability were excluded, as were two autistic children who obtained average results on performance tests. Among the unassessable children there may, of course, have been others of average ability, though examination of the available evidence from the developmental history and partial performance test results leaves this question seriously in doubt in only six cases; it is, however, possible that others of the 22 unassessable children, though below average, were above the intelligence level of most of the group.

The 78 boys and 52 girls were aged 3 to 15 years ($M \times 10.2$ years, $s.d. = 3.46$). Thirty-two children fell below the lower limit of the Revised Stanford-Binet Scale in that they did not succeed with all six of the Year II items, and 22 others had incomplete results or were unassessable. Sixty-nine of the remaining 76 had I.Q.s of under 50 on the Revised Stanford-Binet Scale, the remaining seven ranging up to an I.Q. of 56 on this test. Forty per cent of the group were Mongols; in another 21 per cent of the cases the medical examination revealed definite signs of cerebral abnormality; the rest (39 per cent) showed no clear signs of such abnormality, though five of them were microcephalic (outside -3 s.d. when compared with their age group).

Sixteen of the 130 children had been subjects in the previous investigation; these and two temporary admissions were not included in Part 2 of the enquiry.

3. METHOD

The children were brought by the investigator to her room from the classroom. The observations were made in the course of an investigation of the concepts and manipulative ability of the children. For this they were given various performance tasks, some of which required a demonstration or verbal instructions. Items to test their verbal comprehension were also administered.

A schedule listing the items of behaviour to be observed was used. The items were derived from observations made during clinical work with subnormal children. During the course of modifying this schedule, its consistency was examined when the writer and another examiner* used it for the same 24 children within a period of 3 months. The percentage agreement in classification for the 22 items ranged from 50 to 100 per cent (median 87 per cent). For the item that was the basis of the main classification in this study (response to a social approach) there was a 100 per cent agreement. For three of the other six items used (response to verbal test items; hyperactivity; resistance to demonstration of a task) agreement ranged between 86 and 92 per cent. For two other items (calling the examiner's attention to achievements; and seeking help in difficulty) there was agreement in 75 and 67 per cent of the cases respectively. Some of the disagreement was found to be due to different interpretations of the items rather than to differences in the behaviour of the children, a problem which has been dealt with in subsequent modifications. For the remaining item (the extent of spontaneous vocalizations rated on a 5-point scale) there was 58 per cent agreement. Seven of these disagreements, however, involved a difference of only one point, and only one child was awarded ratings which differed by three points.

Teachers in the training centre were asked to complete a questionnaire on the classroom behaviour of children in their class.

For the second part of the study, data on the previous experience of the children were obtained, after the observations of behaviour had been made, from the case papers, supplemented with additional information from the social workers. For a sample of 30 cases the data were extracted from the case papers by the writer and another psychologist.† The

* Mrs D. J. Stern

† Mrs J. R. Williams.

agreement was so close that it was concluded that this type of information could be reliably extracted by one investigator.

4. RESULTS

(i) *Social behaviour in the individual situation*

The child's social response to the investigator was classified by the presence of absence of the following behaviour, which can be summarized as a response of distress, avoidance or rejection.

<i>Distressed, avoiding or rejecting responses</i>	N
The child cried or whimpered	7
The child initially resisted entering the room	4
The child remained in a corner near the door	3
The child showed marked visual avoidance	13
When spoken to the child hid his face with his arm or began a mannerism which shielded his face	4
The child sat with his head lowered	1
The child ignored the approach	9
The child ignored the approach but initiated contact	5

The last item is intended to characterize the children who themselves approach the investigator (by coming close to her or talking spontaneously) but ignore an approach which is made to them. The last item but one describes children who showed no reaction and stared straight ahead, but did not actively look away when their glance was caught as did those classified as showing visual avoidance. Children showing this behaviour are usually described as 'remote' or 'detached'. All but four of the 14 'ignoring' children manifested in addition one of the other responses. In all, 32 children (Group B) showed one of more of the eight types of response. Changes during the session occurred mainly in the case of the distressed children; only one continued crying the whole time. The visually avoiding children showed the least change.

The other 98 children (Group A) did not show any of the above behaviour most responded positively to a social approach. Group B was on average 2 years younger than Group A. This is probably due to the exclusion from the training centre of severely disturbed children before they reach the leaving age of 16 years. There were no significant sex differences, nor were there any differences between the children who showed signs of cerebral abnormality and those who did not (Mongols excluded), but there were significantly fewer Mongols in Group B (12 per cent) than in Group A (49 per cent) ($\chi^2 = 13.3$, $p < 0.001$, d.f. = 1).

In the previous study the undistressed children were subdivided into the more responsive and the less responsive. Since only 14 children in the present group were characterized as less responsive, this subdivision was not made. Instead, a different one was made on the basis of the ease with which the child's attention could be secured for verbal items, in the individual situation, in four categories as follows:

- (1) Attends well; the child responds at once so that the question does not have to be repeated.
- (2) Attends poorly; some items have to be repeated.
- (3) Difficult to gain and hold the child's attention; repetition of all items is necessary, but the child responds eventually.
- (4) Child's attention cannot be obtained; he makes no response at all.

The number of children in Groups A and B falling in the four categories is shown in Table 1, from which are omitted seven children with extremely limited verbal comprehension.

Table 1. *Social response and attention to verbal test items*

		Child responds to verbal items:				Totals
		(1) At once	(2) After a few repetitions	(3) After many repetitions	(4) Not at all	
Group A	68	18		7	0	93
Group B	3	4		1	22	30
Totals	71	22		8	22	123*

* Five children in Group A and two in Group B with limited verbal comprehension have been omitted from this table.

It will be seen from Table 1 that category 4 (no response) occurs only among the distressed and avoiding children (Group B) and that only three in this group responded at once. Categories 2 and 3 were combined for the test of statistical significance, which indicates a highly significant relation between the type of social response and the degree of attention to verbal items ($\chi^2 = 84.22$, $p. < 0.001$, d.f. = 2).

In subsequent comparisons, Group A has been subdivided into those who attend well (Group A.1) and those in categories 2 or 3 (Group A.2), whose attention can be secured only with difficulty. The five children in Group A with limited verbal comprehension have been placed in Group A.1.

The frequency of occurrence in these groups of five other aspects of behaviour is shown in Table 2, together with the results of the statistical analysis for each of the five items, when the number of children showing each type of behaviour were compared with those not showing it.

Table 2. *Number of children in each social response group showing five types of social behaviour*

	Group A.1 N = 73		Group A.2 N = 25		Group B N = 32		χ^2	$p <$
	N	Per cent of 73	N	Per cent of 25	N	Per cent of 32		
Resist demonstration*	9	16†	15	63‡	20	87§	39.3	0.001
Restless	6	8	10	40	16	50	24.6	0.001
Want to be noticed or respond to praise	56	77	10	40	7	22	30.5	0.001
Seek help in difficulty	20	27	9	36	4	13	4.4	0.1
Vocalize incessantly	5	7	7	44	5	44	13.2	0.01
Vocalize not at all	6	15	4	16	9	28	1.1	0.9

* No information on this item for 16 subjects in Group A.1, one in Group A.2 and nine in Group B.

† Per cent of 57.

‡ Per cent of 24.

§ Per cent of 23.

Table 3. *Social responses and classroom behaviour*

	Group A.1 N = 72 Per cent of 72	Group A.2 N = 25 Per cent of 25	Group B N = 31 Per cent of 31	χ^2	p <
Joins spontaneously in group activities	70	60	10	31.84	0.001
Approaches teacher either frequently or never	19	44	52	12.46	0.01
Indifferent when praised	7	24	29	9.56*	0.01
Ignores commands	4	12	26	7.83*	0.01
Attacks other children	8	32	19	6.76	0.01
Hyperactive	14	32	42	10.27	0.01
Tantrums	18	48	36	9.46	0.01
Tears clothes	10	28	32	8.31	0.02
Throws toys about	15	16	35	5.82	0.1
Manneristic	33	48	55	4.82	0.1

* Group A.2 combined with Group B.

Only 104 children were sufficiently advanced to carry out performance items which required a demonstration. Resistance to this, in varying degrees, took the form of trying to grab the material during the demonstration, or of paying no attention to the demonstration, by not looking at it, or interrupting and talking about something else. This behaviour increased in frequency from Group A.1 to Group A.2 to Group B. The children described as restless in Table 2 kept getting up and wandering about the room, trying to open cupboards and doors and trying to help themselves to other play material. They took no notice when instructed to come and sit down again. A child who lay on the floor and cried and another who refused to sit down at all have been included in the restless group. There were significantly fewer restless children in Group A.1 than in the other two groups.

The third item in Table 2 concerns the steps that children took to ensure that their manipulative activities, and particularly their achievements, were noticed by the investigator, if they thought she was not looking. Some actively drew her attention by touching her arm and pointing to the completed task, or kept looking in order to ensure that she was still watching while they performed the task. Others did not do this, but responded to praise by looking pleased. The rest took no active steps, nor responded to praise. The former behaviour decreased in frequency from Group A.1 to Group A.2 to Group B.

The three groups were not significantly differentiated on the fourth item, attempts by the child to make the investigator complete the task for him when he was in difficulty. The fifth item deals with spontaneous vocalizations, which were rated on a five-point scale. There were only small average differences between the groups, but one or other of the extreme ratings (5: almost incessant talking, or 1: no vocalization at all) occurred more often in Groups A.2 and B than in Group A.1: the figures in Table 2 are the number with ratings of 1 or 5, which were compared with a rating of 2, 3 or 4.

Age trends were examined within each social response group, but few significant

differences were found. In Group A.1 children who actively sought praise were younger (mean age = 9.65 years) than those who responded to praise but did not seek it (mean age = 12.8 years). ($t = 3.71$, $p < 0.001$, d.f. = 54.)

(ii) Social behaviour in the classroom situation

Groups A.1, A.2 and B were compared for their classroom behaviour, with results as in Table 3, which also give the value of χ^2 and the probabilities, when the three groups are compared, for the presence or absence of the items of behaviour listed in column 1. (No information was available in the case of two children who were temporary admissions, not known to the present staff.)

It can be seen that the majority of Groups A.1 and A.2 joined spontaneously in group activities, such as dancing and games; some of Group B could be led into the group and they stayed there for as long as their hands were held, but even this was resisted by 12 of the 31 in this group, who always sat out. Significant differences were found for all items except for mannerisms and throwing toys about. However, manneristic children in Group B tended to show two or more kinds of mannerism, whereas the rest displayed only one kind ($\chi^2 = 3.96$, d.f. 1, $p < 0.05$). The mannerisms reported of children in Groups A.1 and A.2 were mainly of rocking and thumb-sucking. Those reported for Group B included others such as spitting and odd noises, hand waving, jerking of trunk and head or arms, slapping face or hands.

(iii) Social response and early experiences

As mentioned earlier, the number of subjects in this part of the study is 112.

The factors in the child's life before admission to the long-stay hospital which were considered to be adverse are listed below. (The criteria used are those used in the previous study.)

<i>Adverse Factors</i>	<i>Criteria</i>
Home dirty and bare, children dirty and ill-clad	Social workers' statements
Home overcrowded	Parents and more than one child sleeping in the same room
Mother incompetent housekeeper	Father did the housework and catering
Father irresponsible	Gave little money for housekeeping, in arrears with rent (father employed)
Limited finances	Family drawing National Assistance
Chronic physical ill-health in one or both parents	Parent away or incapacitated
Mental ill-health in one or both parents	Psychiatric treatment
One parent deserted the family	—
Severe marital disharmony	Subsequent separation or social workers evidence
Child lived in residential nurseries	—

The 18 children without homes of their own had usually had more than one placement before admission to the long-stay hospital. One child who had lived in one good foster home was included in the not adverse group.

There were 20 conditions, involving 31 children in which the classification was not clear-cut on the basis of the above criteria. In this case the writer and two other psychologists* classified the condition as adverse or not. The other two raters did not know to which social

* Mrs M. D. Cookson and Mr R. R. N. Nation.

response category the children belonged. All three raters placed 11 of the conditions involving 16 children in the not adverse category. In seven of these cases the child was technically illegitimate, but was living with both parents, or with the mother in reasonably good conditions. An example of the other cases is that the home was described as sparsely furnished, but the child was well-cared for. All three raters also agreed in rating one case as adverse in which the mother's illness was not incapacitating but led to her being extremely irritable with the child.

In the remaining 15 cases there was disagreement, and the judgment of the two who agreed was accepted. In eight of these cases the mother was described by the social worker as nervous or anxious, and two of the raters considered this was too vague to be put on a par with cases in which one parent had been a psychiatric referral. Four of the remaining cases were classified as not-adverse by two raters and three were rated as adverse. Whichever category these 15 cases are placed in does not affect the general trend of the results.

Forty-six children had one or more of the adverse factors in their history and 66 had none. The distribution of these in Groups A.1, A.2 and B is shown in Table 4.

Table 4. *Social response and adverse factor in history*

	Group A.1	Group A.2	Group B	Totals
1+ adverse factor	15	12	19	46
No adverse factors	49	11	6	66
Totals	64	23	25	112

Adverse factors occur in 23 per cent of the cases in Group A.1, in 52 per cent of Group A.2 and in 76 per cent of the cases in Group B. These differences are statistically significant at beyond the 0.001 level ($\chi^2 = 21.96$, d.f. = 2).

Group B was 2 years younger on the average than Group A (though Group A.2 was slightly older than Group A.1). (Difficult and unresponsive children tend to be excluded from the training centre before the leaving age of 16 years.) In view of this the incidences of adverse factors in the three social response groups was examined among children who were under 11 years old, and these were similar to those in the whole group: 7 of 24 in Group A.1 (29 per cent); 7 of 9 in Group A.2; and 11 of 17 in Group B (65 per cent). The association is statistically significant in the under-11 group, Group A.2 being combined with Group B in view of the small numbers ($\chi^2 = 8.02$, $p < 0.01$, d.f. = 1).

In view of the lower incidence of Mongols in the distressed and avoiding group, the relation between the social response and adverse factors was examined among the Mongols. Those in the present study were combined with Mongols in the previous study, making a total group of 69. Eight of the 16 in Groups B and A.2 had adverse factors in their history, compared with 10 of the 53 in Group A.1. This difference is statistically significant beyond the 0.01 level (χ^2 with Yates correction applied = 9.23, d.f. = 1).

In the previous study there was a slight but not significant trend for the number of previous admissions to other hospitals or residential nurseries to be related to the type of social response. There was no information on this point in 12 of the present cases. The remaining 100 have been combined with the 90 of the previous study, with results as shown in Table 5. (For those living permanently in nurseries, only hospital admissions were counted.)

It will be seen that there is a trend for Group B (distressed and avoiding) to have been in hospital more often. The differences in Table 5 are significant at beyond the 0.001 level ($\chi^2 = 38.16$, d.f. = 4), the last two columns being combined. Among the 108 children in both studies with no adverse factors in their history, a higher hospital admission rate is significantly more frequent in the combined groups of Groups A.2 and B than in Group A.1 ($\chi^2 = 9.45$, d.f. = 2, $0.01 > p > 0.001$). Of the 32 children from good homes and with no previous hospital admissions, 30 were in Group A.1. There was no relation between the number of previous hospital admissions and the type of social response among those with adverse experiences.

There was a significant trend for Mongols to have had fewer hospital admissions than non-Mongols: in the combined group of 190, 59 per cent of those with no previous hospital

admissions were Mongols, as were 31 per cent of those with one previous admission, and 22 per cent of those with two or more previous admissions ($\chi^2 = 19.16$, $p < 0.001$, d.f. = 2).

The results were examined in order to determine whether any particular experience distinguished Group A.2 from B, but there were no clear trends. Thirty-two per cent of Group B, compared with 17 per cent of Group A.2 had been living permanently away from their homes usually with several placements, but this difference was not statistically significant. The 22 children in Group B who made no response at all to verbal test items probably can be regarded as the most disturbed children in the group. Qualitative examination of the findings suggests that the intensity of the adverse experience rather than the kind may differentiate them from the rest. For example one child (in a materially good home) was kept for most of the day in an empty room because 'he was noisy and dirty', and he was locked in when the mother went shopping. Of the two children from whom the least response could be obtained, one, with six previous hospital admissions, who came from a

Table 5. *Number of hospital admissions and social response, including subjects from previous study*

	Number of hospital admissions				Totals
	0	1	2	3 to 8	
Group A.1	39	34	21	21	115
Group A.2	5	10	1	3	19
Group B	7	21	9	19	56
Totals	51	65	31	43	190

home where there was severe marital disharmony, stood near the investigator, talking all the time, but not listening or responding in any way, not taking any notice of the play material or tasks. The other, with eight previous admissions, whose father was psychotic and whose family lived in a hostel for problem families, lay on the floor and cried the whole time, so that no response at all was obtained. Even in the classroom, if this child notices that he is being observed, he stops what he is doing and covers up his face.

In the previous study the distressed children were older on admission than the undistressed, but there was only a small difference in this respect in the present group. However, the tentative finding of the earlier study that the distressed children from good homes had been in the long-stay hospital for a shorter time than the distressed children from adverse homes was confirmed in the present group.

5. DISCUSSION

Three groups of children, classified on the basis of their social response to an adult, have been distinguished. These groups show significant differences in the individual situation in the relative number who are resistive to demonstration, hyperactive and indifferent to praise, and in the amount of their spontaneous vocalization. In the classroom situation these groups are significantly differentiated by such variables as participation in group activities, initiation of approaches to the teacher, response to praise, hyperactivity and tantrums, being destructive of clothing and ignoring commands.

The types of behaviour of children in Groups A.2 and B are such that the term behaviour disorders might be applied to them. What then has led to this? (Group B was younger on average than Groups A.1 and A.2, but the association was also

significant when children under the age of 11 years were considered separately.) Two conditions which are alleged to give rise to behaviour disorders were common to the whole group, or at least very frequent: all were permanently separated from their mothers, and it is likely that many were suffering from some form of cerebral abnormality. The differences in social response were, however, markedly associated with the presence or absence of adverse circumstances in the child's home before admission or (in a few cases) permanent residence away from home before admission, thus confirming the finding of the previous study on this point. Furthermore, when the groups from both studies were combined, the number of previous hospital admissions was also related to a disturbed social response.

These findings raise the question of the role of cerebral abnormality in the behaviour disorders of the severely subnormal, or at least throw doubt on whether it may be regarded as a single factor. The same point has been made by Harrington and Letemendia (1958) in relation to brain-damaged children of higher intelligence: from the results of their study of such children they concluded that prolonged symptoms of behaviour disorder were more attributable to persistent adverse circumstances in the home rather than to the head injuries. A comparison of severely subnormal children showing clinical signs of cerebral abnormality with those showing no such signs is of dubious value, but even so there was no significant difference when such groups in the present study were compared for the incidence of different types of social response. Cerebral palsy, epilepsy, in fact any form of brain damage, is rare in Mongolism. Although there were relatively fewer Mongols showing behaviour disorder, there were also fewer unfavourable homes and fewer hospital admissions in the Mongol group, and the relation between these environmental factors and the type of social response held when Mongols alone were considered. This result suggests that the smaller incidence of these factors in the Mongol group at least partly accounts for the difference in the incidence of behaviour disorders among the Mongols and the non-Mongols.

In the previous study it was ascertained that the behaviour disorder had been present in most cases from the time of admission onwards. The period between admission and observation ranged up to 8 years in that study ($M = 3.72$ years in the distressed group) and up to $12\frac{1}{2}$ years in the present one ($M = 7.57$ years in Group A.2 and $M = 4.63$ years in Group B). If experience is postulated as a contributory factor in behaviour disorder, this does not imply that the early experience alone contributes to the persistence of the behaviour, or that the effect is assumed to be irreversible. Although there is an increasing amount of evidence in the field of animal experimentation demonstrating the influence of early experience on subsequent behaviour, the question of the irreversibility of the effects is still unsettled (King, 1958). The experimental results of King (1957) with mice suggest that the appropriate social experience may modify the effects of early experiences.

When effects persist, they may be sustained by a continuation of adverse experiences. Indeed, one of the main criticisms of the Bowlby (1951) hypothesis is that the subsequent care of the children in the studies he quotes was far from adequate, a point which is borne out by more recent investigations (Ainsworth, 1962).

In the subnormal group, are there other possible experiential factors which might partly account for the persistence of the behaviour disorder? The experience of admission to a long-stay hospital is one. That children with little cognitive development can react adversely to hospital admission is demonstrated by Schaffer's (1958) observational study of infants. Moreover the subsequent care of the present group in large wards is not the most satisfactory kind. There is also the factor of continued contact with parents. In the previous study it was found that parents in whose home there was an adverse factor visited the child or took him home for the weekend less frequently than did parents with no adverse circumstances in the home. Kellmer-Pringle & Bossio (1960) and Kellmer-Pringle & Clifford (1962), studying intellectually normal children in care, found that contact with parents was the most important factor differentiating severely disturbed from very stable children. Lack of continued contact with parents in the severely subnormal group (particularly with the more advanced children) may also have contributed to the persistence of the behaviour disorder.

The hypothesis that a sequence of unfavourable experiences sustains behaviour disorders raises the question of what other experiences will modify the behaviour. This in turn involves a more detailed consideration of the precise nature and operation of the disturbing experiences, and accounts of this can only be speculative. In the previous paper a tentative hypothesis, which aimed to take account of the limited cognitive development of the children concerned, was put forward (Woodward, 1960), and this was discussed in terms of Hebb's (1949) views on emotional disturbance as a disruption of cortical organization, as for example by the simultaneous arousal of two incompatible cell assemblies. This might occur if an expectancy which had been learned on the basis of previous repeated experience of a situation were aroused by some feature of that situation but was not confirmed by sensory experience, that is, it is unfamiliar. In applying this view to the present problem it is suggested that repeated experience of the unfamiliar might have occurred when the child was being handled, on being fed, washed, etc. It is suggested that the child learns expectancies in relation to tactile, visual and auditory stimuli in the handling situations; if the handling is inconsistent because the mother is harassed by the sort of circumstances that were found in this study, or through changes of people in successive placements, the expectancies are not confirmed, and the disruption which Hebb postulated might be experienced as discomfort or distress. This might again be provoked in subsequent encounters with the unfamiliar. The chances of experiencing the unfamiliar are greatest when the child is removed simultaneously from familiar people and familiar surroundings.

The findings of Douglas & Blomfield (1958) provides some support for this interpretation. In their large follow-up study, symptoms of behaviour disturbance in 5-year-old children were significantly related to separation from the mother if it took the form of the child's going to hospital, but not if the mother went and the child stayed at home. In the latter case, the child was still in familiar surroundings and with familiar people.

If this interpretation proves to be correct, its implication for the care of severely subnormal children in hospital is a reduction in the amount of change the

experience, and a gradual introduction of the unfamiliar, for example the child keeping on his own clothes when first admitted to hospital, the mother staying until the first meal, or bedtime. It would also suggest the hypothesis that the type of family group care already found by Tizard (1960) and Lyle (1960) to make up deficiencies in language and verbal ability, might also have favourable effects on emotional and social behaviour.

REFERENCES

AINSWORTH, MARY D. (1962). The effects of maternal deprivation; a review of the findings and controversy in the context of research strategy. In World Health Organization Public Health Papers, No. 14. *Deprivation of the Maternal Care*. Geneva: W.H.O.

BOWLBY, J. (1951). *Maternal Care and Mental Health*. Monograph Series No. 2. Geneva: W.H.O.

DOUGLAS, J. & BLOMFIELD, J. (1958). *Children Under Five*. London: Allen & Unwin.

HARRINGTON, J. A. & LETEMENDIA, F. J. J. (1958). Persistent psychiatric disorders after head injuries in children. *J. ment. Sci.* 104, 1205-1218.

HEBB, D. O. (1949). *The Organisation of Behaviour*. New York: Wiley.

KELLMER-PRINGLE, M. L. & BOSSIO, V. (1960). Early prolonged separation and emotional maladjustment. *J. Child Psychol. Psychiat.* 1, 37-48.

KELLMER-PRINGLE, M. L. & CLIFFORD, L. (1962). Conditions associated with emotional maladjustment among children in care. *Educ. Res.* 14, 112-123.

KING, J. A. (1957). Relationships between early social experience and adult aggressive behaviour in inbred mice. *J. genet. Psychol.* 90, 151-166.

KING, J. A. (1958). Parameters relevant to determining the effect of early experience upon the adult behaviour of animals. *Psychol. Bull.* 55, 44-58.

LYLE, J. G. (1960). The effect of an institution environment upon the verbal development of imbecile children III. The Brooklands Residential Family Unit. *J. ment. Def. Res.* 4, 14-23.

SCHAFFER, H. R. (1958). Objective observations of personality development in early infancy. *Brit. J. med. Psychol.* 31, 174-183.

TIZARD, J. (1960). The residential care of mentally handicapped children. *Brit. med. J.* 1, 1041.

WOODWARD, M. (1960). Early experiences and later social responses of severely subnormal children. *Brit. J. med. Psychol.* 33, 123-132.

Manuscript received 25 September 1962

Neuroticism, Extraversion and Hypnotizability

BY A. G. HAMMER

University of Sydney

Furneaux and Gibson have claimed that although susceptibility to trance is not correlated with neuroticism or introversion-extraversion, it is associated with particular combinations of scores on those two variables as measured by the Maudsley Personality Inventory. If this is so, it may be inferred that the second order interaction between neuroticism, introversion-extraversion, and score on a hypnotizability inventory known to be correlated with trance susceptibility, will be significant. This is tested in a $3 \times 3 \times 3$ contingency set-up, the inference drawn from the Furneaux-Gibson claim not being supported by the findings. An incidental finding is of a highly significant positive correlation between neuroticism and the hypnotizability questionnaire, but this must be viewed with caution since it was not predicted and could conceivably be accounted for by 'impurities' in the hypnotizability scale.

Furneaux & Gibson (1961) report on the ability of the Maudsley Personality Inventory, which measures traits of neuroticism and introversion-extraversion, to distinguish subjects who will go into trance in a standard induction situation. They found some small tendency for stable subjects to be more hypnotizable than neurotic ones, and no overall association with extraversion. However, from a survey of previous findings they were led to anticipate a more complex relationship, and by way of a χ^2 analysis of the frequency of hypnotizables occurring in each of the four possible combinations of high or low scores on neuroticism and extraversion, came to the conclusions that neurotic extraverts (hysterics?) are *not* susceptible, that neurotic introverts and stable extraverts are, and that stable introverts are about evenly divided. In a later article Furneaux (1961) broadly confirms these findings.

These results are sufficiently at variance with clinical impressions, and with some of the other available experimental findings, to require scrutiny. There are grounds for uncertainty about them. The number of cases involved, and the levels of statistical significance reported, imply the need for replication, especially since in some measure hypotheses were tested via those same data which provoked them. The second study was not an independent confirming one, but an examination of the performance of the same group of subjects on a further variable. In the first study hypnotizability was measured by the achievement of eye catalepsy at the first trial, and in the second by a test of waking ideo-motor responsiveness; the illusory, delusory, compulsive, dissociative and interpersonal features of hypnotic trance have been overlooked. Some of the relationships were investigated by dichotomizing the group at the mean, possibly masking some non-linear associations. The claim clearly needs to be checked.

Full testing of the hypnotic susceptibility of a large group of subjects is a big task, and it was therefore decided to conduct a preliminary study using a questionnaire method of assessing hypnotizability. Previous investigators (e.g. Sutcliffe, 1958; Evans & Thorn, 1962) have shown that a questionnaire consisting of items

such as 'Do you talk in your sleep?' is capable of discriminating somnambules from non-somnambules with considerable success. Although the correlation of the questionnaire with hypnotizability is far from perfect, it may still be inferred that, if the Furneaux-Gibson conclusion is correct, then stable extraverts and neurotic introverts will probably score more highly on a hypnotizability questionnaire than the other groups. The present investigation tests that inference.

METHOD

A questionnaire was prepared which included 24 items from the Maudsley Personality Inventory Neuroticism Scale, 24 from the Extraversion Scale, and the following ten items as a Hypnotizability Scale.

49. Do you draw meaningful 'doodles' without being aware of it while you are drawing?	Yes	No
50. Are you absent-minded?	Yes	No
51. Do you walk, or have you walked in your sleep?	Yes	No
52. Do you talk, or have you talked quite a lot in your sleep?	Yes	No
53. Do you have odd occasions when you cannot remember what has happened over a period of time?	Yes	No
54. Can you lose yourself in a play, or film?	Yes	No
55. Do you sometimes get so lost in reveries that you snap out with something of a surprise?	Yes	No
56. Do you every now and then discover that you have carried out some fairly complicated act without being aware of having done so?	Yes	No
57. Do you have vivid daydreams?	Yes	No
58. Have you sometimes <i>felt</i> that you were acting more as an automaton than as a person with free will?	Yes	No

Although all items were administered together, the three scales were scored separately. Items were scored in the usual '2, 1 or 0' way, so that the range of possible scores for neuroticism was 0-48, for extraversion 0-48, and for hypnotizability 0-20.

The validity of the Hypnotizability Scale is obviously crucial for the design of this investigation. It is attested by the following information about the individual items:

Item 50 discriminates somnambules from non-somnambules ($p < 0.05$; Sutcliffe, 1958).

Items 51 and 52, combined to make either yield a positive indicant, discriminate somnambules from non-somnambules ($p < 0.01$; Sutcliffe, 1958).

Item 53 discriminates somnambules from non-somnambules ($p < 0.05$; Sutcliffe, 1958).

Item 54 discriminates somnambules from non-somnambules ($p < 0.01$; Sutcliffe, 1958).

The combination of the above five items, together with two others not used in the present study, discriminates somnambules from non-somnambules ($p < 0.001$) and loads 0.88 on a factor identified as hypnotizability (Sutcliffe, 1958). Sutcliffe's work was based on a comparison of 24 carefully selected somnambules with 24 controls.

Item 49 correlates 0.22 with a composite hypnosis score, 0.31 with the achievement of amnesia, and 0.33 with achievement of a post-hypnotic suggestion (Evans & Thorn, 1962).

Item 56 correlates 0.35 with a composite hypnosis score (Evans & Thorn, 1962).

An item, worded in a different way but similar to item 55, correlates 0.54 with a composite hypnosis score (Evans & Thorn, 1962).

An item which resembles item 57 except that it refers to 'imagination' instead of 'daydreams' correlates 0.24 with a composite hypnosis score, and 0.44 with achievement of amnesia (Evans & Thorn, 1962).

An item which resembles item 58 except that it refers to 'body movements' instead of 'acts' correlates 0.29 with a composite hypnosis score. (This part of the study makes use of 30 subjects who were carefully selected to cover the full range of hypnotizability as measured by conventional depth scales.) (Evans & Thorn, 1962.)

London, Cooper & Johnson (1962) report a correlation of 0.58 with selected items from the Personal Experiences Questionnaire of Shor (1960), which includes material of the sort

in the present scale. The authors concede that they have capitalized on chance, and that cross-validation is required, but they found it necessary to drop completely only nine of the 44 items. Finally, the corrected split-halves self-correlation for the present set of ten items is 0.59, calculated over 250 cases with the halves matched for 'difficulty'.

Responses were obtained from about 300 subjects. The subjects were of both sexes, and ages ranged from 16 to 60. No selection was deliberately made on grounds of socio-economic status, etc., although ability to understand the meaning of the questions and membership of a 'Western' culture group were required. Subjects were recruited and tested by senior students who no doubt included many of their friends and relatives, so that an educational and socio-economic bias towards the upper end was actually involved. The taking of the questionnaire was in each case individually supervised, but not so closely as to promote in the subjects a need to distort results.

A few forms had to be rejected as 'spoiled'. From the remainder 270 were selected in such a way that on each of the three variables they could be divided into a low, a medium and a high group, each consisting of 90 cases. The frequencies were tabulated in a $3 \times 3 \times 3$ table, and the interrelationships investigated by way of the multi-dimensional χ^2 technique (Lancaster, 1951).

RESULTS

Tables 1, 2 and 3 set out the distributions of the 270 cases on the three variables. It is clear that the subjects have been chosen in such a way that the subgroups

Table 1. *Distribution on neuroticism*

Score	Frequency
0-5	9
6-10	31
11-15	50
16-20	42
21-25	48
26-30	41
31-35	25
36-40	15
41+	9

Table 2. *Distribution on extraversion*

Score	Frequency
0-5	3
6-10	14
11-15	34
16-20	39
21-25	54
26-30	36
31-35	52
36-40	32
41+	6

Table 3. *Distribution on hypnotizability*

Score	Frequency
0-2	53
3-4	37
5-6	56
7-8	34
9-10	46
11-12	20
13+	24

labelled low, medium and high on each of the variables are satisfactory approximations to groups in the general population that would be labelled in the same way. The design of the study makes no further demand on the distributions beyond this one of reasonable spread on the three variables.

Tables 4, 5 and 6 set out the bivariate distributions on neuroticism and extraversion for each grade of hypnotizability separately. These, in effect, set out the

Table 4. *Bivariate distribution on neuroticism and extraversion for high hypnotizables*

		Neuroticism			Total
		Low	Medium	High	
Extraversion	High	6	6	14	26
	Medium	10	6	18	34
	Low	2	10	18	30
Total		18	22	50	90

Table 5. *Bivariate distribution on neuroticism and extraversion for medium hypnotizables*

		Neuroticism			Total
		Low	Medium	High	
Extraversion	High	12	17	6	35
	Medium	9	12	6	27
	Low	4	9	15	28
Total		25	38	27	90

Table 6. *Bivariate distribution on neuroticism and extraversion for low hypnotizables*

		Neuroticism			Total
		Low	Medium	High	
Extraversion	High	17	11	1	29
	Medium	19	8	2	29
	Low	11	11	10	32
Total		47	30	13	90

frequency entries in the $3 \times 3 \times 3$ cells of the three-dimensional contingency table. Tables 7, 8 and 9 give the bivariate distributions between any pair of variables over all cases in the sample, and are thus the marginal surfaces formed by appropriate summings in the $3 \times 3 \times 3$ cube.

The expected frequency in each of the 27 cells is $\frac{1}{3 \times 3 \times 3} \times 270$, i.e. 10, and it is thus possible to compute the total χ^2 over all cells, the degrees of freedom being (27-1). It is also possible to calculate the χ^2 for each of Tables 7, 8 and 9, the

Table 7. *Bivariate distribution on neuroticism and extraversion over all cases*

		Neuroticism			Total
		Low	Medium	High	
Extraversion	High	35	34	21	90
	Medium	38	26	26	90
	Low	17	30	43	90
Total		90	90	90	270

Table 8. Bivariate distribution on neuroticism and hypnotizability over all cases

		Neuroticism			Total
		Low	Medium	High	
Hypnotizability	High	13	27	50	90
	Medium	30	38	22	90
	Low	47	25	18	90
Total		90	90	90	270

Table 9. Bivariate distribution on extraversion and hypnotizability over all cases

		Extraversion			Total
		Low	Medium	High	
Hypnotizability	High	30	34	26	90
	Medium	28	27	35	90
	Low	32	29	29	90
Total		90	90	90	270

expected frequencies in cells being 30 and the degrees of freedom (3-1)(3-1), i.e. 4. Since the three tripartite divisions into 90:90:90 were set arbitrarily there is no point in testing them, but 2 degrees of freedom are lost in each case. Since χ^2 may be partitioned by simple summation, it is possible to calculate the residual χ^2 by subtraction of the χ^2 values for Tables 7, 8 and 9 from the total χ^2 , and this must be evaluated for the degrees of freedom remaining after subtraction of those involved in the marginal surfaces and marginal rows. The results of the computations required are set out in Table 10.

Table 10. Analysis of χ^2 —Summary

Source	χ^2	df	Significance
Total	69.0	26	—
Neuroticism \times Extraversion	18.53	4	$p < 0.001$
Neuroticism \times Hypnotizability	42.8	4	$p < 0.000001$
Extraversion \times Hypnotizability	2.53	4	$p > 0.5$
Residual (Neuroticism \times Extraversion \times Hypnotizability)	5.13	8	$p > 0.5$

DISCUSSION

The χ^2 for the Neuroticism \times Extraversion relationship is clearly highly significant —that is, independence of these dimensions cannot be maintained. The dependence must be allowed for in considering the main issue of this study. Inspection of Table 7 indicates that markedly introverted people tend to be more neurotic than others. The contingency coefficient for these data is about 0.25, but, because of broad grouping and possible curvilinearity, the correlation between the two variables is probably higher than that. This result is not seriously different from those of a number of previous studies.

The χ^2 for the Neuroticism \times Hypnotizability relationship is very highly significant indeed. Inspection of Table 8 reveals that neurotics tend to be hypnotizable, and stable subjects relatively immune. This result is obviously different from that

of Furneaux and Gibson, but the issue is one over which there has been in the past some controversy. For the moment it is sufficient to say that it must be allowed for in considering the main issue of the study.

The χ^2 for the Extraversion \times Hypnotizability relationship is not significant. This is in line with previous studies, although it is of course contrary to the view often supported by clinicians. There is some reason to doubt whether hysterics actually do score as extreme extraverts, on the MPI, and it is possible that the scale, despite its name, does not measure what clinicians have thought of as extraversion. In any case, when we reflect on the cases of extreme spontaneous hypnosis—that is, of dual personalities such as Mrs Beauchamp and Eve White—it is doubtful whether the overt personalities of somnambules should be expected to show up as 'extraverted' in the Eysenck sense.

Our main interest is in the residual χ^2 . It is this value which tells us whether, after we have made allowances for the main effects just discussed, there are particular combinations of neuroticism and extraversion scores which are associated with unexpectedly high or low frequencies of hypnotizability—that is, it is this value which tests the Furneaux-Gibson claim. The obtained value is clearly insignificant and the claim that hypnotizability is associated with particular combinations of score on the two MPI variables is therefore rejected. The rejection is less than certain since hypnotizability was not assessed directly, but by way of a variable known to be associated with it. However, it may well be that the questionnaire is not much worse an indication of trance susceptibility than a single test of eye catalepsy following a brief first induction procedure.

Because of some lingering doubts about the validity of the hypnotizability scale, and because there are grounds for believing that rather less than 30 per cent of the population are fully hypnotizable, the results on the other two tests of the 20 cases who scored 14 or more on the hypnotizability questionnaire were scrutinized. Table 11 sets these results out. Obviously there are no grounds here for doubting the earlier findings.

Table 11. *Bivariate distribution on neuroticism and extraversion of 20 highest scorers on hypnotizability*

		Neuroticism			Total
		Low	Medium	High	
Extraversion	High	1	0	4	5
	Medium	0	0	7	7
	Low	1	1	6	8
Total		2	1	17	20

Whilst the present results call the Furneaux-Gibson view into serious doubt, it must be added the positive finding of an association of hypnotizability with neuroticism should be viewed with great caution. True, there is nothing improbable in the notion that proneness to dissociative phenomena predisposes an individual to develop neurotic defences against anxiety under conditions of stress, and the sometimes encountered immunity of neurotic patients to trance might well

disappear if sufficient care and time were taken to allay their initial interfering anxieties. On the other hand, it must be conceded that whilst the items of the hypnotizability scale may be saturated with a dissociation factor, they may additionally be saturated with a neuroticism factor orthogonal to it. It is not unlikely that these items pick out people who have a history of dissociative experiences, and that they also pick out people who are ready to complain of their felt inferiorities and discontents. Although each of the ten items correlates with the neuroticism score (the data are available from the author), the contingency coefficient derived from the previous χ^2 calculation between 'hypnotizability' and neuroticism is only 0.37, so that there is plenty of elbow room for the questionnaire to correlate with each of two independent factors (neuroticism and dissociation). It remains necessary, therefore, to test the unpredicted association of hypnotizability with neuroticism by way of thorough actual trance induction.

REFERENCES

EVANS, F. J. & THORN, W. A. F. (1962). The prediction of hypnotic phenomena from questionnaire items. Paper submitted for publication. *J. clin. exp. Hypnos.*

FURNEAUX, W. D. (1961). Neuroticism, extraversion and suggestibility. *J. clin. exp. Hypnos.* 9, 195-214.

FURNEAUX, W. D. & GIBSON, H. B. (1961). The Maudsley Personality Inventory as a predictor of susceptibility to hypnosis. *J. clin. exp. Hypnos.* 9, 167-177.

LANCASTER, H. O. (1951). Complex contingency tables treated by the partition of χ^2 . *J. R. statist. Soc. Ser. B*, 13, 242-249.

LONDON, P., COOPER, L. M. & JOHNSON, H. J. (1962). Subject characteristics in hypnosis research, II and III. *J. clin. exp. Hypnos.* 10, 13-21.

SHOR, R. E. (1960). The frequency of naturally occurring hypnotic-like experiences in the normal college population. *J. clin. exp. Hypnos.* 8, 151-163.

SUTCLIFFE, J. P. (1958). Hypnotic Behaviour: Fantasy or Simulation? Unpublished thesis, Fisher Library, University of Sydney.

Manuscript received 8 October 1962

Perceptual Sensitivity and Defense in Amputee Children

By LOUISE CENTERS

Private Practice, Los Angeles

AND RICHARD CENTERS

University of California, Los Angeles

In order to test the hypothesis that amputee children will exhibit more conflict and defensiveness in the form of delayed recognition of an amputee figure which resembles their own body image than will non-amputee children, a tachistoscopic speed-of-recognition test was given to two groups of children. Amputee children manifested a significantly longer reaction time in recognizing any difference between two figures presented tachistoscopically, one of which had an arm missing. When they did recognize a difference, however, they more readily attributed it to a difference in the arms of the figures than did the non-amputee children. The latter more frequently accounted for the difference in the figures by referring to other parts of the bodies or proposing that the figures were posed differently. The findings are interpreted as supporting the hypothesis of greater defensiveness of the amputee child, with a difference in perceptual sensitivity or vigilance also being involved.

While whole volumes have been devoted to the psychological problems concerning handicapped children who are deaf, blind, mentally retarded or deficient, work having specific reference to the amputee child is conspicuous by its absence. Research on amputation has been confined largely to empirical investigations of effects of such injuries to adult war veterans. Although it is granted that the number of children malformed from birth or from diseases or injuries sustained in their early years is small in comparison to the number of other handicapped children, this factor is probably only one of several which may have inhibited investigation of psychological problems of the former. An interesting possibility is that the limbless body is itself such an anxiety-provoking object and so threatening to the bodily integrity of the would-be researcher that he finds his attention readily preoccupied elsewhere.

For whatever reason, questions as to the amputees' psychological reactions to his malformation are important not only for what light answers to them may throw on the psychological and social adjustment of such children, but also for what information may be obtained with regard to psychological processes of more general and universal import.

The most directly relevant prior research with respect to the experiment reported here is that of Noble, Price & Gilder (1954) who investigated the effects of the perception of amputation by adult male war veterans with such afflictions. Their findings were that denial was the most prominent defense, but with displacement and projection also manifested.

PROBLEM

Since the amputee child, like adult amputees, may represent a threat to the integrity of the body image of both himself and those in his environment, it was hypothesized that he would manifest more conflict and defensiveness about his body image than would the non-amputee child about his own image. More specifically, it was predicted that the threshold for recognition or a stimulus figure involving amputation would be higher for the amputee child than for the non-amputee child.

METHOD

A total of 52 subjects participated in the experiment. The amputee sample consisted of 26 Caucasian children, 16 boys and 10 girls, between 5 and 12 years of age with upper extremity amputations or anomalies, who were patients of the Child Amputee Prosthetics Project of the University of California at Los Angeles.

Twenty-two of the 26 children had congenital amputations or anomalies, while 4 had sustained traumatic amputations before the age of 3. The types of amputations ranged from unilateral transcarpals through shoulder disarticulations with a majority, 17 cases, having amputations below elbow and wrist. All were otherwise in good health, had normal vision without correction and at least average intelligence as measured by the Columbia Mental Maturity Scale.

The non-amputee control group consisted of 26 children who were matched with the amputee group for chronological age, sex and approximate mental age on the CMM Scale. Socio-economic status distribution of the two groups was similar, being lower middle class.

Three drawings were presented to each subject by means of a modified Dodge tachistoscope in familiar physical surroundings. One was the drawing of a tree, the other two were drawings of either a boy or a girl. The ones presented depended on the sex of the subject. The two figures of the boy or girl were identical except that one had the right arm missing with only the sleeve showing and the other had both arms present.

After the subject was seated in front of the tachistoscope and was focusing on the fixation point, he was then given the following instructions: 'If you keep looking at the "X" you may see some other things. Watch very closely and tell me everything you see.' Figure 1 of the tree was then exposed for 0.01 second and at successively decreasing speeds until the subject identified the tree.

After this threshold value for the tree was established the test series was begun. Figures 1, 2 and 3 (male figures) were presented to boys and figures 1, 4 and 5 (female figures) to all girls in random order at speeds of 0.01, 0.02, and then 0.02 of a second intervals through 0.2 of a second, then at 0.10 second intervals through 0.81 seconds, the upper limit of the tachistoscope. When either of the two figures was recognized as a boy or girl, the subject was asked to describe it in detail. If the subject had not identified any difference between the two human figures at 0.81 seconds, then the cards were presented successively with constant illumination until the child mentioned a difference.

RESULTS

For each subject five threshold values were established:

1. Threshold for Tree was the minimal speed at which the subject identified the stimulus as a tree. This value represented a threshold value against which the others could be compared and also in the pre-test situation served to familiarize the subject with the apparatus and the task.

2. Threshold for Person was the minimal speed at which the subject identified either the amputee figure or the non-amputee figure in terms of its sex.

3. Threshold of First Difference was the minimal speed at which the subject verbalized any difference between amputee and non-amputee figures. For example: 'He's smiling and the other one wasn't.'

4. Threshold of Real Difference was the minimal speed at which the subject verbalized a difference in the arms of the two figures for any reason. For example: 'Can't see one arm; he has his arms behind him and the other has his down to the sides.' In some cases this value coincided with the Threshold of First Difference.

5. Verbalization Time was the minimal speed at which the subject verbalized the absence of one of the arms on 3 or 5.

One comparison of the amputee and non-amputee groups was made by computing difference scores for each matched pair of subjects on these five thresholds. Another comparison was made by computing difference scores for each pair between several thresholds.

The Wilcoxon Matched-Pairs Signed-Ranks Test, a non-parametric statistic, was applied to those Threshold values where magnitude of differences between subjects was within the measurable limits of the tachistoscope.

A difference in the predicted direction was revealed between the two groups in Threshold of First Difference that was significant at the 0.01 level of confidence. In other words, the amputee children took longer to verbalize any difference between the two figures than the non-amputees. No significant differences were found between groups in the Thresholds for the Tree, for the Person, or for the Threshold of Real Difference.

The *T* of the Wilcoxon test for the difference between the Threshold of Real Difference and the Threshold of First Difference was found to be significant beyond the 0.05 level of confidence. This means that once the amputees recognized any difference between the two figures they more readily attributed it to the amputation. However, these same amputees took a significantly longer period than the non-amputees between recognition of a human figure and noticing any difference between the two figures. This difference was significant at the 0.005 level of confidence.

The Sign test was applied to the Verbalization Time data since the magnitude of the differences between paired scores could not be determined in several cases. Nineteen of the 26 non-amputees took longer to verbalize the idea of amputation than did their amputee partners. In fact, ten of the non-amputees never said the arm was missing, but accounted for its absence in some other way, such as its being behind the back, even when the stimulus was exposed under constant illumination. Only five amputees accounted for it in any way other than being missing. This difference is significant at the 0.03 level of confidence. It might be argued that this finding could be accounted for in terms of the non-amputees' unfamiliarity with amputation. However, several of the non-amputees had in their classes a child with an upper extremity amputation.

DISCUSSION

Other tests (Centers, 1958) have established that amputee children have a realistic awareness of their body images and how they differ from non-amputees. Thus, it could be expected that they would recognize a figure which was a rough approximation of the body image as far as sex and site of amputation is concerned as readily

as non-amputee children recognize a representation of their body images. That is, one could expect this if perception of the body image were only influenced by sensory stimulation.

Results of the Recognition Test demonstrate, however, that there is something which interferes with the rapidity of perception of a difference between figures for the amputee children and delays it beyond that of non-amputee children.

Our interpretation is that conflict and defensiveness in the form of denial of the threatening stimulus can account for the delay in recognition of any difference between the figures. This conclusion is strengthened by the rather neat agreement we find with Noble, Price & Gilder (1954) who in their studies found that denial was the most prominent defence employed by adult war-amputees in figure drawings. Linn (1953), also, has shown that a visual stimulus presented tachistoscopically will be recognized less rapidly if it conflicts with the wishes and interests of the subject. It is as if for the amputee child to recognize the non-amputee figure constitutes no threat, probably because he is aware that most people have intact bodies and, more importantly, he has probably internalized the positive value of the culture for having all limbs. Recognition of amputation, and consequently the differences in the figures, however, arouses negative feelings, conflict, and anxieties which he has about his own body.

Contrasting the delay in recognition of any difference, it was found that the amputees more quickly recognized the Real Difference as one in the arms once they realized there was a difference in the figures at all. Spence (1957) has shown that both denial and perceptual vigilance can operate in the same individual. This suggests that this finding can be interpreted as a manifestation of vigilance or sensitivity, for the amputees are more likely to relate this difference to the area in which they are most sensitive—where their body images differ from non-amputees.

REFERENCES

CENTERS, LOUISE (1958). Perceptual and social effect of amputation in children. Unpublished doctoral dissertation, University of Southern California.

LINN, L. (1953). The role of perception in the mechanism of denial. *J. Amer. Psychol. analyt. Assn.* 1, 690-705.

NOBLE, D., PRICE, D. B. & GILDER, R. (1954). Psychiatric disturbances following amputation. *Amer. J. Psychiat.* 110, 609-613.

SPENCE, D. (1957). A new look at vigilance and defense. *J. abnorm. soc. Psychol.* 54, 103-108.

Manuscript received 23 October 1962

A Cross-cultural Investigation of the Relationship Between Introversion-Extraversion, Social Attitudes and Anti-social Behaviour

BY ARON WOLFE SIEGMAN

University of Maryland School of Medicine

This paper describes a number of studies which were undertaken in order to test Eysenck's theory that the introversion-extraversion factor is a significant source of variance in (1) authoritarian or 'tough-minded' attitudes, (2) religiosity, and (3) anti-social behaviour.

There was no support for the hypothesized positive correlation between extraversion and tough-minded or authoritarian attitudes. Introversion-extraversion was a significant source of variance in subjects' religious behaviour but the direction of the relationship varied with religious background (Protestant and Jewish) and sex. Finally, the predicted positive correlation between extraversion and anti-social behaviour was confirmed in a group of male students, but not in a comparable female group. This finding was interpreted in terms of a sex-role identification theory of anti-social behaviour.

The studies to be described in this paper were undertaken in order to test Professor Eysenck's theory of individual differences in social and political attitudes as well as his theory of anti-social behaviour. In *The Psychology of Politics* (1954) Eysenck concludes that a substantial part of individual differences in social and political attitudes is a result of variability in relation to a conservatism-liberalism factor and a tough-mindedness-tender-mindedness factor. Moreover, Eysenck suggests that the tough-mindedness-tender-mindedness dimension is an expression in the attitudinal field of the extraversion-introversion factor, which he has found to be a significant source of variance in many other areas of human behaviour. Thus, this theory, like the theory proposed by Adorno, Frenkel-Brunswik, Levinson and Sanford in the *Authoritarian Personality* (1950), attempts to relate individual differences in the attitudinal field to basic personality variables. As a matter of fact Eysenck maintains that the F-scale which was developed by Adorno *et al.* (1950) as a measure of authoritarian ideology is essentially an index of tough-mindedness.

The link between tough-mindedness-tender-mindedness and extraversion-introversion is explained in the following manner: (1) it is assumed that the socialization of man's aggressive and sexual impulses proceeds on the basis of conditioning, (2) a number of studies have demonstrated that introverts condition more easily than extraverts, (3) consequently introverts should be more thoroughly socialized than extraverts, and (4) this will be reflected in pacifism and sexual restraint, attitudes which define the tender-mindedness end of the tough-mindedness-tender-mindedness factor. In making the last mentioned of the four arguments, Eysenck has the following to say: 'If the reader will glance back at the attitudes characterizing the two poles of the T-factor, he will notice that attitudes characterizing tough-mindedness are essentially concerned with immediate satisfaction of aggressive (war, persecution, hanging, flogging, etc.) and sexual (birth

control, easier divorce, more abortion, etc.) impulses. Tender-minded attitudes, on the other hand, are concerned with ethical and religious ideas which act as barriers to such satisfactions and which, since time immemorial, have been part of the socialization process' (1954, p. 259). This quotation has implications for the psychology of religious behaviour by suggesting a relationship between tender-mindedness and religion and by implication between introversion and religiosity. More importantly, this theory has implications for the very dynamics of the socialization process. In Eysenck's own words: 'According to this view, socialized behaviour in the adult has as its basis anxiety and fear responses to anti-social acts of an overtly aggressive or sexual character; these responses are conditioned in childhood and cohere together according to the principles of stimulus generalization (aided by verbal identification). The resulting barrier to the immediate satisfaction of every passing impulse (conscience, "inner light", super-ego) is stronger in introverts than in extraverts, due to the greater strength of the conditioning process in the former. Taken to extremes, this gives us the neurotic introvert, the dysthymic, who is over-socialized and prone to phobias and anxieties due to his over-strong conditioning equipment, and the neurotic extravert, the psychopath, who is under-socialized and prone to anti-social acts due to his defective conditioning equipment. This conditioned ethico-religious barrier to impulse satisfaction in the introvert also emerges in the attitude field in the form of "tender-minded" attitudes, while the relative absence of such barriers gives rise to "tough-minded" attitudes' (1960, p. 15). Clearly, the extravert can be expected to exhibit relatively more acting out of anti-social impulses than the introvert.

Although the theory has been spelled out in great detail it is yet to be empirically validated. Fortunately, the assumptions postulated by this theory readily lend themselves to empirical verification. The studies to be reported were performed over a period of years, and had as their objective the testing of these assumptions and their implications.

STUDY I

The major objective of this study was to evaluate that part of Eysenck's theory which relates the extraversion-introversion factor to the tough-minded or authoritarian attitude syndrome and to certain specific attitudes which, according to Eysenck, characterize this syndrome.

Statement of hypotheses

1. There is a negative correlation between extraversion and tender-minded attitudes (as measured by Eysenck's Tender-mindedness (T) scale), and a positive correlation between extraversion and authoritarian ideology (as measured by the F-scale).
2. In the factor analytic study which provided the basis for Eysenck's 'two factor' theory of social attitudes, the items dealing with minority prejudice and discrimination obtained some of the highest loadings on the tender-mindedness-tough-mindedness factor (Eysenck, 1954, pp. 129-130). This is not surprising if we conceive of this factor as a measure of the ability to control aggressive and hostile

impulses. Consequently, we could expect a positive correlation between extraversion and ethnic prejudice.

3. As is evident from the earlier quoted excerpts, Eysenck assumes that religious institutions are major socializing agents, and that religiosity is an index of the degree to which one has become socialized. According to this assumption we could expect a negative correlation between religiosity and extraversion.

Subjects

The subjects of this study were 43 male undergraduate students at the University of North Carolina. All subjects in this group were members of the dominant white Protestant sub-culture.

Procedure

All Ss were administered the following questionnaires: Eysenck's Tendermindedness (T) scale (the ten T₁ items in Eysenck, 1954, pp. 277-279), Eysenck's Extraversion (E) scale (Eysenck, 1956), a 29-item F-scale (Adorno *et al.*, 1950, pp. 255-257), the six items from the Ethnocentrism (Eth) scale which pertain to Negroes (Adorno *et al.*, 1950, p. 142) and a religious attitude scale (RAS) developed by the author (Siegman, 1962d).

Results and discussion

Table 1 lists Ss' mean and S.D. scores on the various questionnaires, and Table 2 lists the intercorrelations between the attitude and personality measures. As can be seen in Table 2, none of the predicted relationships between extraversion and

Table 1. *Mean and standard deviation scores obtained by Ss in Study 1 on personality and attitude measures*

Measures	(N = 43)	
	Mean	S.D.
E-scale	31.26	6.65
T-scale	7.40	2.22
F-scale	98.65	15.21
Eth-scale	19.49	6.89
RAS-scale	65.26	12.70

attitudes were confirmed by the data of the present study.‡ Furthermore, the results of the present study also fail to confirm Eysenck's claim that both the T-scale and the F-scale measure the same, or at least a very similar attitude syndrome.

Table 2. *Intercorrelations (r) of personality and attitude measures in Study 1*

Measures	T-scale	E-scale
E-scale	0.07	
F-scale	0.07	0.21
Eth-scale	-0.10	0.09
RAS-scale	0.46†	0.30*

In Tables 2-6, * indicates $p < 0.05$,
 † $p < 0.01$, and italics $p < 0.10$.

‡ A subsequent study by Eysenck (1961) has, on the other hand, found the predicted correlation between E and T scales.

The only finding of this study which is consistent with Eysenck's general position is the significant positive correlation between the religiosity and tender-mindedness scales ($r = 0.46$, $p < 0.01$). This correlation, however, may be a result of the fact that the T-scale contains a number of items referring to religious beliefs. Thus, the positive correlation between the religiosity and tender-mindedness scales is probably artifactual and spurious. This conclusion is supported by the fact that in the present study religiosity was associated with authoritarian ideology ($r = 0.53$, $p < 0.01$), which according to Eysenck reflects a tough-minded orientation. The religious items which are contained in the tender-mindedness scale may also be responsible for the failure to obtain the expected negative correlation between the T and F scales, since in the present group religiosity is apparently related with an authoritarian and tough-minded rather than a tender-minded orientation.

The finding that religiosity is associated with an authoritarian ideology, although in contradiction with Eysenck's results in which the religiosity items obtained substantial loadings on the tender-mindedness factor, is consistent with the results obtained in a number of other studies (Adorno *et al.*, 1950; Gregory, 1957; Levinson, 1954; Levinson and Huffman, 1954). One can think of two potential explanations for the contradiction between these two sets of data. Already in the *Authoritarian Personality* (Adorno *et al.*, 1950) it is pointed out that only a fundamentalist or conservative religious outlook was found to be associated with an authoritarian ideology, and that a liberal religious philosophy was not associated with authoritarian attitudes. Evidence in support of this distinction has also been found in another study (Siegman, 1962d). Perhaps this distinction can also provide the basis for a resolution of the contradiction between the results of the present study and those obtained by Eysenck in relation to religiosity and authoritarian ideology. Perhaps the index of religiosity which was used in the present study is a measure of conservative religiosity and hence its association with authoritarian attitudes. Eysenck's items, on the other hand, may refer to a more liberal religious orientation and hence their association with 'tender-minded' attitudes. However, a careful inspection of the T-scale items concerned with religion failed to support this interpretation. Another possible explanation of the conflicting results between the two sets of data is in terms of the considerable social pressure which exists in the United States in favour of some form of religious behaviour. One of the outstanding features of the authoritarian personality is his need to conform (Adorno *et al.*, 1950). Thus in a society in which there is considerable social pressure in favour of some form of religious affiliation and observance, it is to be expected that the authoritarian will conform in this respect also. In societies in which there is no pressure in favour of religious conformity, however, the religious individual may tend to hold, as found by Eysenck, tender-minded or non-authoritarian attitudes. (Again, all this is based on the assumption that an authoritarian ideology is equivalent to a tough-minded outlook, and a non- or anti-authoritarian ideology is equivalent to a tender-minded orientation.) The fact, however, that religious belief was found to be associated with an authoritarian ideology in a group of Israeli college students (Siegman, 1962d) who live in a culture in which there is, if anything, social pressure against religious affiliation and observance—would cast

some doubt on this second distinction. It should be pointed out, however, that despite the non-religious orientation of a large segment of the Israeli population, the groups which are religiously oriented do exert considerable social pressure in favour of religious conformity, so that the distinction may still hold. Whatever the explanation for the contradiction between the results obtained by Eysenck and those of the present and other studies may be, the important point is that the relationship between religious attitudes and other attitudinal clusters is apparently not invariant, but a function of culture and perhaps other factors.

Another finding of the present study which is in the opposite direction from that suggested by Eysenck, is the significant positive correlation between the RAS-scale and the E-scale. Perhaps this finding, too, can be explained in terms of the considerable social pressure in favour of some form of religious affiliation and observance which characterizes the United States social scene. Eysenck himself has suggested that extraverts are more sociable and therefore also more likely to succumb to social pressure (Eysenck, 1957, p. 24). Thus, extraverts would be more religious where the social pressure is in this direction. Whatever the validity of the hypothesis that extraverts are more sociable, the E-scale is certainly a measure of social contact needs. The following items illustrate the point: 'Would you be very unhappy if you were prevented from making numerous social contacts?', 'Do you like to mix socially with people?', 'Are you inclined to limit your acquaintances to a select few?', 'Do you usually take the initiative in making new friends?' and 'Do you like to have many social engagements?' Consequently, the finding of a positive correlation between the E-scale and the RAS.

STUDY 2

The major concern of this study was again that part of Eysenck's theorizing which relates certain attitudes, such as authoritarian ideology, ethnocentrism and religiosity to the introversion-extraversion dimension. Although the results of the first study were essentially negative, it will be recalled that some of the findings were related to special features of the United States culture. Consequently, it was felt that additional testing of Eysenck's theory under different cultural conditions was necessary before conclusions about its validity can be made. The present study represents an attempt at such further investigation.

Subjects

The subjects of this study were 40 undergraduate students at Bar-Ilan University, Israel. All *Ss* in this group were of European and middle-class background.

Procedure

Ss were administered Eysenck's Extraversion (E) scale, a ten-item F-scale specially adapted for the population of this study (Siegman, 1961b), and two ethnic prejudice scales, one measuring prejudice against the local Arab population (AP-scale) and the other measuring prejudice against Jewish North African immigrants (NP-scale) (Siegman, 1961b). Thirty-two *Ss* in the group also took Thurstone and Chave's (1929) Attitude Toward the Reality of God (RBS) scale, a measure of religious belief, and Foa's (1948) Sabbath Observance (ROS) scale, an index of religious observance. According to Eysenck, introversion should be associated with less authoritarian attitudes, less prejudice and greater religiosity.

Results and discussion

As can be seen in Table 3, there was no significant correlation between the E-scale and the measure of authoritarian ideology. Although significant correlations were obtained between the E-scale and the indices of ethnic prejudice, they

Table 3. *Correlations (r) between the E-scale and attitude measures in Study 2*

Attitude measures	(N = 40)	E-scale
F-scale		-0.08
AP-scale		-0.29
NP-scale		-0.34*
RBS-scale		-0.20
ROS-scale		-0.34*

are in the opposite direction from that which Eysenck's theory would predict. This finding is rather surprising and the reason for it is not immediately apparent. Perhaps this finding is also related to the point which has been made earlier, that the E-scale is a measure of social conformity. Since in Israel there is considerable social disapproval of prejudiced attitudes, the high E-scale scorer, being a social conformist, would be reluctant to admit such attitudes, and hence the negative correlation between the E-scale and the measures of ethnic prejudice. This explanation, however, is admittedly post-hoc speculation, and subject to further investigation.

The correlations between extraversion and the two measures of religiosity are in the expected direction, although only one—the negative correlation between the E-scale and Foa's Sabbath Observance scale—is significant. These findings, including the fact that the significant negative correlation was between extraversion and religious observance and not between extraversion and religious belief, confirm Eysenck's theoretical formulation. According to Eysenck there should be a negative correlation between extraversion and religiosity, because extraverts are more difficult to socialize and because religiosity reflects successful socialization. This rationale is clearly relevant to the observance dimension of religiosity, and less so, if at all, to the faith or belief dimension. Thus, the results of the present study are clearly consistent with Eysenck's theoretical formulation of the relationship between the introversion-extraversion dimension and religiosity.

Elsewhere (Siegmund, 1962a) the author also hypothesized a positive correlation between introversion and religiosity, but on the basis of a different rationale from that of Eysenck. It is suggested that the religious attitude, the belief in a supernatural power which stands in some kind of relationship to man and the physical universe, tends to attribute to natural phenomena more than the immediately given, and that such a belief is likely to be held by individuals with a relatively active imaginative life. Furthermore, it is assumed that introverts tend to be more imaginatively inclined, and that extraverts are more objectively oriented. This assumption is supported among others by the finding that introverts produce more *M* responses on the Rorschach than do extraverts (Carrigan, 1960). This rationale implies a negative correlation between extraversion and religious observance. Thus,

unlike the results of a previous study in which different and less sensitive measures of religious belief and observance were used (Siegman, 1962a), the data of the present study are consistent with Eysenck's rather than the author's theoretical formulation.

STUDY 3

The purpose of this study was to obtain further data on the relationship between the introversion-extraversion dimension and religiosity. It will be recalled that in Study 1, which used United States male college students of the Protestant faith, extraversion was found to correlate *positively* with a general index of religiosity. In Study 2, which used Israeli college students of the Jewish faith, the opposite relationship was obtained: extraversion correlated *negatively* with indices of religious belief and observance. One obvious explanation for the differential relationship between introversion-extraversion and religiosity in the two groups is that they represent two different religious backgrounds (Jewish and Protestant). The two groups, however, also differ in relation to other variables, such as cultural background and sex composition. In the present study, therefore, an attempt was made to investigate the relationship between the introversion-extraversion dimension and religiosity in more directly comparable Protestant and Jewish groups.

Subjects

Three different groups of subjects participated in this study. The first group consisted of 54 Protestant and 24 Jewish male second year medical students at the University of Maryland. The second group consisted of 31 Protestant and 13 Jewish, male, first year students at the University of Maryland School of Medicine. The Protestant and Jewish *Ss* were of comparable socio-economic background. A third group was added in order to evaluate the possible effect of sex on the relationship between introversion-extraversion and religiosity. This group consisted of 55 Protestant female third and fourth year students at the University of Maryland School of Nursing. These *Ss* were similar in cultural and socio-economic background to the male Protestant groups.

Procedure

Subjects in group I were administered Eysenck's E-scale and a ten-point religiosity self-rating scale (RSR-scale). This self-rating scale has been found to correlate 0.86 (d.f. = 35) with the earlier described RAS-scale (Siegman, 1962d).

Subjects in groups II and III were administered Eysenck's E-scale, a Guttman-type scale developed by Goldsen *et al.* (1960) to measure religious belief, and a Guttman-type scale measuring church or synagogue attendance (religious observance). Table 4 lists *Ss*' scores on these last two measures.

Table 4. *Religious belief and observance scores of different groups in Study 3*

Groups	N	Goldsen Scale		Church (or Synagogue) attendance scale	
		Mean	S.D.	Mean	S.D.
Protestant males	31	3.17	1.20	3.07	1.65
Protestant females	55	3.74	0.78	3.50	1.16
Jewish males	13	2.92	1.12	1.92	1.04

Results

The results of the present study confirm the implication of the previous findings that the relationship between introversion-extraversion and religiosity varies with

religious denominational background. Furthermore, the data of the present study suggest that the relationship between introversion-extraversion and religiosity also varies with sex. As can be seen in Table 5, in the Jewish groups the introverts tend to be more religious, but among Protestant males the opposite seems to be the case. Furthermore, whereas among Protestant males the extraverts tend to be more religious, in the female Protestant group the trend is in the opposite direction. Although the negative correlations between the E-scale and the two measures of religiosity in the female group were not significant, they did differ significantly ($p < 0.05$) from the positive correlations between the same measures in the Protestant male group.

Table 5. *Correlations (r) between the E-scale and religiosity indices in Study 3*

Groups	Sex	N	E and Goldsen Scale	E and Church attendance	E and RSR Scale
Group I (Protestants)	M	54			0.35†
Group I (Jews)	M	24			-0.34
Group II (Protestants)	M	31	0.35*	0.46†	
Group II (Jews)	M	13	-0.67†	-0.52*	
Group III (Protestants)	F	55	-0.16	0.05	

The fact that among Protestant males the extraverts tend to be more religious whereas among Protestant females the tendency is in the opposite direction, may be related to Argyle's (1959, p. 79) theory that religious behaviour in women is motivated by internal processes while men are religious as a result of external sanctions. Stated somewhat differently, the evidence suggests that personal conviction plays a relatively stronger role in female religiosity, and social conformity plays a relatively stronger role in male religiosity. If this thesis is correct, and if the E-scale really is, as it had been suggested, a measure of social conformity, we could indeed expect a positive correlation between the E-scale and religiosity in male but not in female groups. Furthermore, if we assume that there is stronger social pressure in favour of religious affiliation and observance among United States Protestants than among United States Jews, the above arguments would also explain the positive correlation between the E-scale and religiosity among Protestant males but not among Jewish males. That there actually is relatively more social pressure in favour of some form of religious observance among Protestants than among Jews has been argued by Herberg (1960), and may be reflected in the finding of the present study that Protestants attend religious services more frequently than do Jews ($t = 2.88; p < 0.05$) although the two groups do not seem to differ in the intensity of their religious beliefs (see Table 5). The opposite relationship can be observed in relation to sex. The female Protestant group obtained significantly higher belief ($t = 2.28, p < 0.05$) but not higher observance scores than the male Protestant group. Assuming that the belief scale taps personal convictions and that the observance scale is relatively more sensitive to external pressures, these findings confirm the thesis that female religiosity is more of the personal rather than the social conformity kind. These explanations for the differential relationships between introversion-extraversion and religiosity are

admittedly speculative and require further confirmation. The important point, however, is that the relationship between introversion-extraversion and religiosity is not invariant, but rather a function of such variables as denominational background and sex, which is not accounted for by Eysenck's theoretical formulation.

STUDY 4

Ultimately, Eysenck's hypotheses about the relationship between various specific attitudes and values and the introversion-extraversion dimension are based on the more general hypothesis—which is derived from a set of carefully spelled out postulates (Eysenck, 1954, 1960)—that with all other factors being equal it is easier to socialize introverts than extraverts. Proceeding from this general hypothesis he then relates various attitudes which are assumed to reflect successful socialization to the introversion-extraversion dimension. Thus, while the general hypothesis has implications for the psychology of attitudes and values, it is even more directly relevant to the psychology of anti-social behaviour. As pointed out by Eysenck, the hypothesis that aggressive and sexual impulses are more readily socialized in introverts than in extraverts, implies that the introversion-extraversion dimension is a significant source of variance in crime and delinquency. Bartholomew (1961) attempted to test this hypothesis in an empirical manner by comparing the E-scale scores of first offenders, recidivists and controls. The only significant difference was between first offenders and recidivists, the latter group obtaining the higher mean score. Although this finding is consistent with Eysenck's theoretical position, the fact that no significant differences were obtained between first offenders and controls, nor between recidivists and controls is clearly contrary to Eysenck's theory. Elsewhere (Siegman, 1961a) the author argued that since many of the E-scale items refer to social participation, the prisoners who are restricted in their social activities may have obtained a spuriously low E-scale score. This difficulty can be avoided by the use of behavioural indices of introversion-extraversion rather than the E-scale in comparing offenders and non-offenders. In carrying out such studies, the author found that young offenders did, in fact, obtain significantly higher scores, in the extraverted direction, on the following three behavioural correlates of the introversion-extraversion dimension: I.Q./vocabulary ratio, perceptual rigidity and time judgment (Siegman, 1961a).

The assumption that officially adjudicated criminals are representative of law breaking individuals in general has been vigorously challenged by Short & Nye (1957, 1958). These authors have pointed out that social class and other background factors seriously prejudice one's chances of being arrested and sentenced. According to these authors a much better way to investigate criminal behaviour is to ask people, under conditions of complete anonymity, how frequently they have engaged in various kinds of legally prohibited activities. It is then possible to correlate such admissions with whatever variables one wishes to investigate. By using this method one can control such potentially confounding factors as socio-economic background, etc. As a matter of fact, this method was used by Kinsey and his associates (1948) in their study of sexual behaviour. Short and Nye, in their

studies, have demonstrated that with adequate assurances of confidentiality, people are willing to confess their transgressions in other areas as well. On the basis of such admissions these authors succeeded in constructing a Guttman-type scale consisting of a variety of legally and/or socially disapproved activities. Although this scale has adequate reliability (Short & Nye, 1958), questions have been raised about its validity. Thus, it has been suggested that the scale may not be so much a measure of anti-social behaviour as of the readiness to admit socially unacceptable behaviour (Siegman, 1961a, 1962b). In a recent study, however, no significant correlations were obtained between admitted anti-social behaviour scores and Christie's Social Desirability scale (Siegman, 1963). These findings indicate that under conditions of complete anonymity *Ss* apparently do give relatively honest replies, and do not attempt to present themselves in a more acceptable light. In the present study such anonymous admissions of anti-social behaviour will be used to evaluate Eysenck's hypothesized relationship between introversion-extraversion and anti-social behaviour.

Subjects

The subjects of this study were 44 male first year students at the University of Maryland School of Medicine, and 70 female third and fourth year students at the University of Maryland School of Nursing.

Procedure

All *Ss* were administered Eysenck's E-scale and an anti-social behaviour (ASB) scale. The ASB-scale which was used in the present study is a Likert-type scale, and consists of all those items in the original Short & Nye (1958) scale which correlated significantly, in a group of college students, with *Ss*' total score (Siegman, 1961a, 1962b). A determination was also made of *Ss*' socio-economic background on the basis of criteria developed by Meier & Bell (1959). *Ss* were requested not to indicate their names, in order to ensure full anonymity.

Results and discussion

The male group obtained significantly higher ASB-scale scores (Table 6), which is consistent with previous findings (Short & Nye, 1958). The two groups did not

Table 6. *ASB-scale scores and socio-economic status scores of male and female Ss in Study 4*

Groups‡	N	ASB-scale		Socio-economic status	
		Mean	S.D.	Mean	S.D.
Male	44	7.55	3.40	19.53	4.90
Female	70	4.92	2.71	19.10	4.12

‡ The only significant difference in this table is in relation to the ASB-scale ($t = 4.24$, $p < 0.01$).

differ significantly in relation to socio-economic background, nor was this variable a significant source of variance in *Ss*' ASB-scale scores. A similar finding has been interpreted by Short & Nye (1958) as evidence in favour of the conclusion that there is no relationship between socio-economic background and criminality. The fact that a disproportionate number of officially adjudicated criminals come from

the lower socio-economic levels they attribute to the selective nature of this criterion. An alternative interpretation will be suggested in the course of this discussion.

In the present study the extraversion and anti-social behaviour scales correlated positively in the male group ($r = 0.35$, $p < 0.05$), and negatively in the female group ($r = -0.22$, $p < 0.10$). The positive correlation between the E-scale and anti-social behaviour in the male group is consistent with Eysenck's theoretical formulation. Yet, if we are to explain this finding in terms of Eysenck's hypothesis that extraverts are more difficult to socialize, how is one to account for the fact that female extraverts seem to be *more* socialized? Moreover, in another but similar group of male students the author (Siegman, 1961a) failed to obtain significant correlations between Ss' ASB-scale scores and their scores on two behavioural indices of introversion-extraversion (time estimation and vocabulary/I.Q. ratio). These findings suggest that introversion-extraversion is not related to anti-social behaviour as measured by the ASB-scale. But how is one to reconcile this conclusion with the earlier cited findings suggesting that officially adjudicated criminals are more extraverted than comparable controls? One possibility is that there is no causal relationship between introversion-extraversion and criminal or anti-social behaviour as such, and that extraverts are merely more likely to be arrested, either because of the way in which they commit their crimes or for some other reason. This explanation, however, is rather speculative and not very plausible. Another possibility is that introversion-extraversion is, in fact, causally related to officially adjudicated criminality, but not to the criminal and anti-social activities which are sampled by the ASB-scale. This explanation assumes that there are different categories of criminal activities, in the sense that they have different causal antecedents. Such differences may be related to the type of crime (theft, assault), chronicity, etc. This formulation is in opposition to the point of view that criminality is a continuous variable, with major and chronic criminality defining one of its poles, impeccable probity defining the other pole and occasional petty criminality the midpoint. Such a conceptualization of anti-social behaviour has been advocated by Gough (1961) and by Short & Nye (1958), and is held implicitly by many others. To conceive of anti-social behaviour as a continuous variable has obvious theoretical advantages. The fact, however, that many variables which are associated with officially adjudicated crime (e.g. impulse control (Siegman, 1961a), introversion-extraversion, n achievement (Siegman, 1961c), etc.) are not associated with admitted criminality as measured by the ASB-scale suggest that there is more than one kind of criminal behaviour. If so, socio-economic background (it will be recalled that this variable was not associated with Ss' ASB-scale scores) may be another variable which is differentially related to different kinds of anti-social behaviour.

Elsewhere (Siegman, 1962c, 1962e) the author has spelled out and cited evidence in support of a role identification theory of anti-social behaviour (of the kind measured by the ASB-scale rather than major crimes). This theory, together with the assumption that the E-scale is a measure of social conformity, may provide an explanation for the differential relationship between the E-scale and anti-social

behaviour (as measured by the ASB-scale) in males and females which was found in the present study. The theory states that because our culture condones aggressive and acting out behaviour in boys but frowns upon such behaviour in girls, identification with the male role will be associated with greater anti-social behaviour, and identification with the female role will be associated with less anti-social acting out. Thus, males who identify with their sex role—i.e. the male role—will evidence greater anti-social behaviour, whereas females who identify with their sex role—i.e. the female role—will evidence less anti-social behaviour. Assuming that the E-scale is an index of social conformity—and by implication a measure of conformity with one's sex role—a high E-scale score should indeed be associated with a greater degree of anti-social acting out in males and the opposite in females.

As a general comment on the various studies reported in the present paper one may point out that although the relationships between introversion-extraversion—as measured by the E-scale—and other variables frequently failed to confirm Eysenck's predictions, it was always possible to explain the data assuming that the E-scale is a measure of sociability. Although sociability is *one* of the traits which characterize the extravert (Eysenck, 1957, 1961), it does not define—in the Eysenckian scheme—the introversion-extraversion dimension. A definition of introversion-extraversion *primarily* in terms of socialization would deviate from Eysenck's conceptualization of this dimension in terms of cortical excitation-inhibition balance. If it is correct that the E-scale is primarily a sociability index, and that sociability is only one among many of the extravert's traits, we may have an explanation for the discrepancies between results obtained with the E-scale and those obtained with other criteria of introversion-extraversion (e.g. introversion-extraversion was found to be a significant source of variance in conditionability when the former is defined in terms of psychiatric diagnostic categories, but not when defined in terms of the E-scale (Holland, 1960). To avoid this, as well as other problems inherent in the use of questionnaire type tests, it may be best to use a variety of indices of the introversion-extraversion variable.

Study I, which was conducted during 1956-1957 while the author was at the University of North Carolina, was supported by a grant from the University's Behavioral Sciences Research Fund.

Study II, which was conducted during 1959-1960 while the author was at Bar-Ilan University, was aided by a grant from the Committee for the Encouragement of Research of the Israeli Labor Organization.

Study III was conducted in 1960-1961, Study IV in 1961-1962, at the University of Maryland.

The author acknowledges his gratitude to Jacob Jonah and Tsadok Baruch of Bar-Ilan University for their help in the collection of the data, to Professor Eysenck for making the E-scale available before its publication, and to Professor Christie for making available his revised and as yet unpublished Social Desirability Scale.

REFERENCES

ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. J. & SANFORD, R. N. (1950). *The Authoritarian Personality*, 1st ed. New York: Harper.

ARGYLE, M. (1959). *Religious Behavior*. Glencoe, Ill.: Free Press.

BARTHOLOMEW, A. A. (1959). Extraversion-introversion and neuroticism in first offenders and recidivists. *Brit. J. Delinq.* 10, 120.

CARRIGAN, P. M. (1960). Extraversion-introversion as a dimension of personality. *Psychol. Bull.* **57**, 329-360.

EYSENCK, H. J. (1954). *Psychology of Politics*. London: Routledge and Kegan Paul.

EYSENCK, H. J. (1957). *Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.

EYSENCK, H. J. (1956). The questionnaire measurement of neuroticism and extraversion. *Riv. Psychol.* **50**, 113-140.

EYSENCK, H. J. (1960). Levels of personality, constitutional factors, and social influences: An experimental approach. *Int. J. soc. Psychiat.* **6**, 12-24.

EYSENCK, H. J. (1961). Personality and social attitudes. *J. soc. Psychol.* **53**, 243-248.

FOA, U. G. (1954). An equal interval scale for the measurement of attitudes towards Sabbath observance. *J. soc. Psychol.* **27**, 273-276.

GOLDSEN, R. K., ROSENBERG, M., WILLIAMS, R. M. JR. & SUCHMAN, E. A. (1960). *What College Students Think*. Princeton, N.J.: Van Nostrand.

GOUGH, H. G. (1960). Theory and measurement of socialization. *J. consult. Psychol.* **24**, 23-30.

GREGORY, W. E. (1957). The orthodoxy of the authoritarian personality. *J. soc. Psychol.* **45**, 217-232.

HERBERG, W. (1960). *Protestant, Catholic, Jew*. Garden City, N.Y.: Anchor.

HOLLAND, H. L. (1960). Measures of learning and conditioning. In *Experiments in Personality* (ed. H. J. Eysenck). London: Routledge and Kegan Paul.

KINSEY, A. C. (1948). *Sexual Behavior in the Human Male*. Philadelphia: W. B. Saunders.

LEVINSON, D. J. (1954). The intergroup relations workshop: Its psychological aims and effects. *J. Psychol.* **38**, 103-126.

MEIER, D. L. & BELL, W. (1959). Anomia and differential access to the achievement of life goals. *Amer. Soc. Rev.* **24**, 189-202.

NYE, F. I. & SHORT, J. F. (1957). Scaling delinquent behavior. *Amer. Soc. Rev.* **22**, 326-331.

RIESMAN, D. (1950). *The Lonely Crowd: A Study of the Changing American Character*. New Haven: Yale University Press.

SHORT, J. F. & NYE, F. I. (1958). Criterion of deviant behavior. *Soc. Problems*, **5**, 207-214.

SIEGMAN, A. W. (1961a). Theories of juvenile delinquency: Some empirical investigations. Symposium paper read at XIVth Intl. Cong. Appl. Psychol. Copenhagen, Denmark.

SIEGMAN, A. W. (1961b). A cross-cultural investigation of the relationship between ethnic prejudice, authoritarian ideology, and personality. *J. abnorm. soc. Psychol.* **63**, 654-655.

SIEGMAN, A. W. (1961c). A comparison of young offenders and a control group in relation to *n* achievement. *J. clin. Psychol.* **17**, 379-380.

SIEGMAN, A. W. (1962a). Personality and socio-cultural variables associated with religious behavior. *Arch. fur Religionpsychologie*, **7**, 96-104.

SIEGMAN, A. W. (1962b). Personality variables associated with admitted criminal behavior. *J. consult. Psychol.* **26**, 199.

SIEGMAN, A. W. (1962c). Theories of anti-social behavior: Some empirical data. Paper presented at annual meetings, East. Psychol. Ass., April, 1962.

SIEGMAN, A. W. (1962d). A cross-cultural investigation of the relationship between religiosity, ethnic prejudice and authoritarianism. *Psychol. Rep.* **11**, 419-427.

SIEGMAN, A. W. (1962e). Some social and psychological factors associated with juvenile delinquency and anti-social behavior. Paper presented at Vth World Cong. Soc., September, 1962, Washington, D.C.

SIEGMAN, A. W. (1963). Father identification, emotional maladjustment and anti-social behavior in male college students. Submitted for publication.

THURSTONE, L. L. & CHAVE, E. J. (1929). *The Measurement of Attitudes*. Chicago: Chicago University Press.

Inter-generation Attitudes

By F. MUSGROVE

Department of Education, University of Leeds

A survey of adolescents' attitudes to adults and of adults' attitudes to 'teenagers' was carried out in 1962 in two socially contrasted Midland areas. An open-ended sentence-completion schedule was the principal method used with both adolescents and adults.

Both boys and girls made a much higher proportion of favourable references to adults than the cross-section of the adult population made to them. The proportion of favourable attitudes to adults did not decrease markedly as adolescents advanced in age, nor did the proportion of favourable attitudes to peers increase, though the proportion choosing adults as preferred companions dropped rapidly between 11 and 13. The behaviour and attributes ascribed to mothers were favourably commented on by a particularly high proportion of both boys and girls.

No marked social-class or area differences in attitude were apparent. Neither boys nor girls distinguished sharply between the roles of mothers and fathers. Attitudes of boys to fathers and of girls to mothers showed slight signs of deterioration soon after puberty.

INTRODUCTION

The purpose of the inquiry reported below was to examine the attitudes towards adults of adolescent boys and girls in the secondary schools of two socially contrasted areas in the north Midlands, and the attitudes of adults in the same area towards 'teenagers'. Although little in the way of systematic inquiry into the attitudes of representative samples of adults towards adolescents was known to the author, it was expected that the survey might show generally hostile, critical and negative attitudes, but that there might be social-class, and perhaps sex, differences. Since this inquiry was completed the work of Eppel (1962) has been published showing the very guarded attitudes of a highly selected group of English adults (London magistrates, probation officers and youth leaders) towards the present generation of teenagers.

American social scientists like Riesman (1950) and Mead (1948) have alleged 'peer-group solidarity', conformity and loyalty among adolescents and the progressive rejection, after later childhood, of parental standards, values, guidance and companionship. Psychologists like Stanley Hall (1905) and Jersild (1957) have reported empirical research which shows American children after the age of 11 abandoning parents as models for their future careers and as advisers when they are in difficulties.

The wide currency of these ideas led the author to suppose, when he began his inquiry, that his adolescent subjects would show a far greater preference than younger children for the society of their coevals; that with advancing years they would make in an attitude test an increasing proportion of favourable and approving references to their coevals and a declining proportion to adults in general and parents in particular. Psychoanalytical theory and empirical inquiries like Liccione's (1955) into harmony and tension in child-parent relationships suggested

the probability that at or soon after puberty boys' attitudes to fathers and girls' attitudes to mothers would be particularly hostile and critical.

The need to check the truth of these views, and particularly their applicability to English society, whatever their validity might be in America, is ever more apparent as studies in England and France throw doubt on the simplicity of this thesis regarding the parent and peer-group relationships of adolescents. The work of Morris (1958) in England has questioned the notion of peer-group solidarity, conformity and loyalty (perhaps particularly among the middle-class young). He found that the proportion of adolescents who felt that they should support their peers in certain situations of conflict with adult authority 'declined sharply with age'. (On the other hand, the proportion who thought they would, in fact, support their friends remained relatively constant.) The French bourgeois family, according to Pitts (1960), still enfolds its adolescents more completely and securely than the American family: 'The peer group cannot expect the type of loyalty that the American or English (?) peer group may claim.' Ariès (1960) offers a similar analysis: '*la famille moderne se retranche du monde, et oppose à la société le groupe solitaire des parents et des enfants.*' Even in America the loyalty to peer-group standards and conformity to peer-group behaviour are now being seriously questioned. Murphy (1947) reported studies of 'cordiality' at an American boys' summer camp which suggested that high status was determined less by conformity than by self-sufficiency; Lucas & Horrocks (1960), in an attempt to isolate the psychological needs of adolescents, failed to distinguish a specific need for peer-group conformity; and Riley, Riley & Moore (1961) have recently found American adolescents less 'other-directed' than they had supposed, perceiving the different expectations of parents and peers, and themselves preferring patterns of behaviour which fell somewhere in between.

The thesis that adolescents become increasingly critical of adults and transfer their approval to their coevals seemed to the author to merit further investigation. The inquiry was also designed to reveal qualitative as well as quantitative changes in attitude. In particular it was hoped to throw some light in an English setting on the view of Parsons & Bales (1956) that in the modern nuclear family the role of the mother is 'expressive' and of the father 'task-oriented' and 'instrumental'. Whether the subjects in this inquiry made this particular distinction or not, it was expected that the qualities ascribed to parents would change between later childhood and adolescence, and that the behaviour of mothers and of fathers would be sharply distinguished. The attitudes to parents which have been reported in America were not expected with any confidence. Bronson, Katten & Livson (1959) concluded from a study of American family life that both sons and daughters perceived mothers more often than fathers as exercising strong authority in the home; Nimkoff (1942) has argued that the parent who habitually exercises discipline will not be preferred to the parent who does not. The (probably) different authority structure of the English family made it seem likely that the subjects in the present inquiry, while they might ascribe to mother an expressive and integrative role, would not see her as the main agency of family discipline, or the object of such strong hostility as Liccione found in his studies of American adolescent girls.

METHODS

The present inquiry was carried out in the schools and among the adult population of two socially contrasted areas of the Midlands which were chosen by the author because of the contrast they presented, because they were of manageable size, and because the head teachers of the schools were interested in the project and co-operative.* One town which was chosen is situated in the north Midlands, has a population of 51,000 and is engaged in heavy industry, textiles and mining. The second area, 60 miles away, is a suburbanized village 10 miles from a large city. The original agricultural village is now heavily overlaid with new and expensive owner-occupied houses: professional and business men live there and work in the nearby city. Professional and managerial families predominate and the total population of the village is now 1800. Most of the children of secondary school age attend a new junior comprehensive school, still comparatively small, in its early stages of growth, which also serves other predominantly middle-class neighbourhoods on the city's outskirts.

Three questionnaires were used with a selected population of schoolchildren in the two areas: (1) asking for preferred companions in a variety of activities and situations, (2) a social distance scale, and (3) an open-ended sentence-completion test. The children were asked to indicate on a list who would be their preferred companions in ordinary leisure activities: 'If you were going on an outing one Saturday or in the holidays (for instance, to the pictures, to a football match, or for a picnic) and could take any two people with you, whom would you most like to take? Father, mother, some other grown-up person, another, a friend as old as you of the same sex, another, a friend as old as you of the opposite sex, another, a friend older than you of the same sex, another, a friend older than you of the opposite sex, another, a friend younger than you of the same sex, another, a friend younger than you of the opposite sex, another.'

In order to discover how choice of companions might be influenced by the difficulties involved in the situation envisaged, the questionnaire also asked for preferred associates in three situations of ascending order of difficulty: 'If you were having a party and could invite any five people, would you want all of them to be over 30 years of age, some of them, or none? If you were going on a long and difficult journey with five other people, would you wish them all to be over 30, some of them or none? If you were in a position of great danger and could get five people to come to your aid, would you want all five to be over 30, some of them, or none?'

Social distance scales were also devised to measure the 'social distance' between the generations. No previous attempt to use social distance scales for this purpose was known to the author, but it seemed reasonable to him to adapt methods employed by Bogardus (1947) and reported by Banton (1959) in measuring social distance between racial groups, and by Shaw (1954) for measuring the social distance between children of the same race and age. The social distance scales in Bogardus, Banton and Shaw were inspected for items which might be rephrased to apply to different age groups: in Banton and Bogardus seven items indicating different degrees of social distance were used, in Shaw five. Twenty statements which seemed to indicate varying degrees of social acceptance of other age groups were drawn up by the author. Although the 'interval' between these statements was clear from inspection, it was decided to confirm this impression by asking 12 adult members of a university extra-mural class to sort these statements out on a five-point scale ranging from complete social acceptance to extreme social rejection. From each of the five categories a statement was taken which had been unanimously assigned to it by members of the class. Social distance scales were then devised to show attitudes to adults (over 30), to old people (over 65), and to people of the subjects' own age. The following questionnaire was used, and others appropriately modified to apply to different age groups:

'Tick one statement with which you agree:

1. I should like to spend all my spare time with people of my own age and with no-one else.
2. I should like to spend a good deal of my spare time with people of my own age, but not all of it.

* I am indebted to Miss Dorothy Wass for her assistance in administering and analysing the questionnaires.

3. I don't mind people of my own age being around, but don't want much to do with them.
4. I don't mind having people of my own age around just once in a while.
5. I like it best when people of my own age aren't around at all.'

In order to discover changing attitudes to their coevals, to adults in general, and to parents in particular, with increasing age, the adolescents in this inquiry were also given an open-ended, sentence-completion test. Riley *et al.* and Morris forced responses to particular situations: the former presented their subjects with personality 'vignettes', the latter with descriptions of problem situations in which there was conflict between the attitudes, values and behaviour of friends, on the one hand, and parents or other adult authorities, on the other. In the author's inquiry as free a response as possible was sought, although the problems of subsequent analysis were thus considerably increased. All the subjects were simply asked to complete in any way they liked 12 sentences for which openings were provided: 'Mothers are. . . . Mothers can. . . . Mothers should. . . . Fathers are. . . . Fathers can. . . . Fathers should. . . . Grown-ups are. . . . Grown-ups can. . . . Grown-ups should. . . . Boys (or girls) of my age are. . . . Boys (or girls) of my age can. . . . Boys (or girls) of my age should. . . .' The age, sex, and father's occupation (as the best single criterion of 'social class') were obtained for each child who completed a questionnaire. All questionnaires were completed anonymously.

A questionnaire was sent by post to a random sample of adults in the two selected areas. Every two-hundred-and-fiftieth person was selected from the electoral roll of Midland town, and every tenth person from the electoral roll of suburbanized village. Again the intention was to permit the freest possible response. The Eppels inquired only among adults in positions of authority and asked for answers to four specific questions (with two subsidiaries); for instance, 'Do you think that the morals of young people today differ from those that applied when you were in your teens?' In the author's inquiry six brief sentence openings were provided and the respondents were asked to complete them in any way they wished which gave their true opinion of the general run of teenagers as they had experienced them. The six sentence openings were: 'Teenage boys are. . . . Teenage boys can. . . . Teenage boys should. . . . Teenage girls are. . . . Teenage girls can. . . . Teenage girls should. . . .' The questionnaires were completed anonymously, but details of age, marital status, family composition, occupation or former occupation if retired (and husband's occupation in the case of married women) were asked for.

RESULTS

The sentence-completion schedule which had been prepared for the use of adolescents was filled in by 302 boys and girls aged 11 to 15 (inclusive) in two secondary modern schools in Midland town and by 54 boys and girls aged 11-13 years in the junior comprehensive school which serves the suburbanized village. These numbers were obtained by taking a random sample of one-third from the registers of each year of each school. The choice-of-companions questionnaire was completed by 200 additional children aged 9 and 10 in two junior schools in Midland town and by a further 224 boys and girls in a third secondary modern school. The social distance scale was completed by 72 of the junior schoolchildren and by 133 of the 14- and 15-year-old boys and girls in the secondary schools. The children of Midland town were predominantly (77 per cent) from the homes of manual workers; only 46 per cent of the children of the comprehensive school were from families of this social-occupational level.

(a) *Age of preferred companions*

Seven hundred and seventy-eight boys and girls between the age of 9 and 15 indicated the two persons on the prepared list whom they would most like to have

with them on an outing to the pictures, a football match, or a picnic. The percentage of choices given to parents declined significantly with each year of age except between the age of 10 and 11. The percentages given to parents by boys and by girls were not significantly different except at age 13, when girls gave fewer choices than did boys (2.1 per cent compared with 13.5; $p < 0.01$) and similarly at age 14 (4.1 per cent compared with 10.3; $p < 0.01$). By age 15 boys had fallen to the same low level of parent-preference as girls—2.5 per cent in both cases.

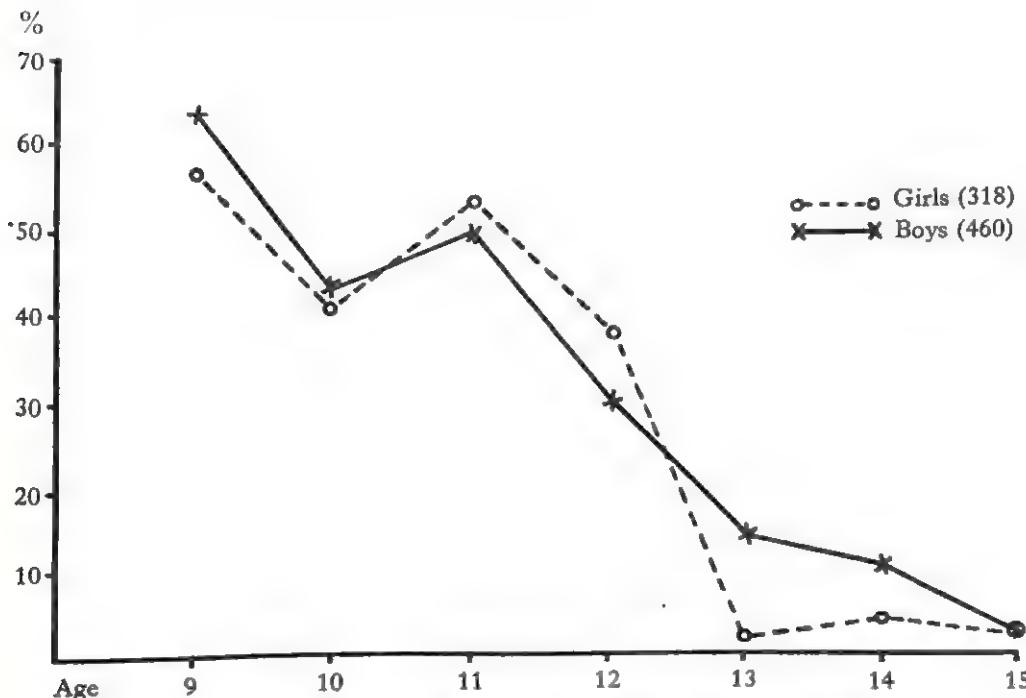


Fig. 1. Percentage of companionship choices given to parents.

There was no tendency at any age for independence of parents to be correlated with level of intelligence or with social-class background. Intelligence scores (N.F.E.R. Group Verbal Tests) were obtained for two classes of boys aged 9 and 10: 60 per cent of the boys of below average intelligence gave at least one choice to their parents, only 47 per cent of the boys of above-average intelligence did so. The correlation with intelligence was not significant ($r_{bis} = 0.15$, S.E. 0.1). Twenty-eight girls of the same age showed no significant tendency to choose parents more often the lower their intelligence ($r_{bis} = 0.264$, S.E. 0.1). The correlation between intelligence and parent-preference among 44 boys aged 14 and 15 was not significant ($r_{bis} = -0.21$, S.E. 0.5). Children from the homes of manual workers gave the same proportion of their choices to parents as children from the homes of non-manual workers.

There was a significant tendency among both boys and girls of all ages for preference for adult companionship to be correlated with the difficulty-level of the task involved. Seventy boys and girls aged 9 and 10, and 132 aged 14 and 15,

were asked whom they would prefer to have with them in three situations of ascending order of difficulty (Table 1).

Table 1. *Preference for adults in three situations*

	Age: 9 and 10			Girls (32)			
	Boys (38)	All	Some	None	All	Some	None
Party	—	—	20	18	1	22	9
Journey	3	—	32	3	4	26	2
Danger	11	—	26	1	20	12	—
		C = 0.54			C = 0.55		
Age: 14 and 15							
	Boys (94)			Girls (38)			
	Boys (94)	All	Some	None	All	Some	None
Party	—	—	29	65	—	8	30
Journey	3	—	67	24	—	28	10
Danger	35	—	56	3	10	25	3
		C = 0.59			C = 0.55		

The social distance scales, which were completed by 205 children, gave no support to a view which is today commonly expressed, that the young are 'nearer to' the elderly than to the middle-aged; but showed that while young children (age 9 and 10) place no greater social distance between themselves and adults than they do between one another, young adolescents (age 14 and 15) hold adults at a significantly greater social distance than their coevals. Seventy per cent of the 9- and 10-year-old boys expressed a strong preference (by endorsing the first two statements on the social distance scale) for the society of over-thirties, 90 per cent for the society of their own age-group. This difference is not statistically significant. Eighty-eight per cent of the 9- and 10-year-old girls expressed a strong preference both for over-thirties and for their own age-mates. At age 14 and 15, on the other hand, only 36.8 per cent of the 95 boys expressed a strong preference for over-thirties, 96.8 per cent for their coevals ($\chi^2 = 77$, d.f. = 1, $p < 0.001$). The corresponding figures for 38 14- and 15-year-old girls were 34.2 per cent and 94.7 per cent ($\chi^2 = 30$, d.f. = 1, $p < 0.001$). The proportion of boys and girls expressing a strong preference for the society of over-sixty-fives did not differ from the proportion expressing a strong preference for the society of over-thirties.

(b) *Adolescents' self-pictures*

Three hundred and fifty-six boys and girls in three secondary schools completed the open-ended sentence-completion schedule. From the statements made about boys and girls of their own age it was possible to build up the self-picture which prevailed in each age-sex group. Thus in completing the sentence: 'Boys (or girls) of my age can . . .' defects were mentioned ('be spiteful'), physical competence ('climb trees', 'play hockey'), technical competence ('help', 'cook'), and social maturity ('go out with boys'). Boys and girls aged 11 to 13 (predominantly pre-pubescent) are compared under these four headings with boys and girls aged 14 and 15 (predominantly post-pubescent) (Table 2).

Table 2. Adolescents' self-pictures ('Boys (or girls) of my age can . . .')

	Defects	Physical competence	Technical competence	Social maturity
Boys (N 56) age 11-13	21.5% (12)	41.1% (23)	10.7% (6)	26.8% (15)
Boys (N 116) age 14-15	39.6% (46)	10.3% (12)	7.8% (9)	42.2% (49)
Diff.	$p < 0.05$	$p < 0.001$	NS.	$p < 0.05$
Girls (N 60) age 11-13	31.6% (19)	8.3% (5)	13.4% (8)	46.6% (28)
Girls (N 102) age 14-15	32.3% (33)	5.9% (6)	2.9% (3)	58.8% (60)
Diff.	NS.	NS.	$p < 0.01$	NS.

The younger girls mentioned activities which denote social maturity more often than the younger boys (c.R. 2.64, $p < 0.01$). Younger girls differ little from older girls in their over-all self-picture, except that they claim technical competence more often. The most striking contrast is between younger and older boys: the latter mention defects, but also claim social maturity, more often, and physical competence significantly less frequently.

In completing the sentence beginning: 'Boys (or girls) of my age should . . .' younger girls again differed little from older girls in the type and frequency of their responses; both expressed a particularly strong sense of personal responsibility (see

Table 3. *Behaviour prescribed by adolescents for adolescents ('Boys (or girls) of my age should ...')*

1	2	3	4	5
Respect for adults	Independence of adults	Friendliness to age-mates	Sense of responsibility	Physical skills
'Consider parents' point of view' etc.	'Be able to stay out till 11 p.m.', 'Be allowed more freedom',	'Mix together', 'Go around together'	'Be sensible', 'Act their age'	'Do sport', 'Keep fit'
			Girls (N 58, age 11-13)	
17.2% (10)	8.6% (5)	12.1% (7)	56.9% (33)	5.2% (3)
			(N 104, age 14-15)	
15.4% (16)	7.7% (8)	19.2% (20)	52.9% (55)	5.8% (6)
Diff. NS.	NS.	NS.	NS.	NS.
			Boys (N 57, age 11-13)	
8.8% (5)	10.5% (6)	17.6% (10)	43.8% (25)	19.3% (11)
			(N 112, age 14-15)	
13.4% (15)	8.0% (9)	10.7% (12)	61.7% (69)	6.2% (7)
Diff. NS.	NS.	NS.	$p < 0.05$	$p < 0.01$

Table 3, col. 4). Younger boys differed markedly from older boys in their more frequent mention of physical skills and their less frequent references to social-personal responsibility. Statements could be classified into the five categories indicated in Table 3.

(c) *Attitude to adults*

In completing the sentences which began: 'Mothers can . . ., Fathers can . . .' the exercise of authority was mentioned ('can hit you if they like', 'can be too bossy'), but more frequently at all ages personal qualities unconnected with the exercise of authority ('can be generous', 'can get worried') and technical proficiency ('can mend things', 'can dig the garden'). The younger boys and girls were alike in the frequency with which their responses fell into these three categories; the older girls mentioned personal attributes of both fathers and mothers more often than boys of the same age, and technical proficiency less.

Neither the younger nor the older boys and girls in this inquiry appeared to see mothers and fathers in the contrasted 'expressive' and 'instrumental' roles postulated by Parsons: the younger children did not mention technical competence more often when referring to fathers than to mothers. In the older age groups, however, girls saw both mothers and fathers as predominantly 'expressive', boys saw both mothers and fathers as predominantly 'task-oriented' and 'instrumental'. There was no significant tendency at any age for either boys or girls to mention the exercise of authority by one parent more often than they mentioned its exercise by the other (Table 4).

Table 4. *Adolescents' perceptions of parents ('Mothers can . . . fathers can . . .')*

	Mothers			Fathers		
	Exercise of authority	Personal attributes	Technical proficiency	Exercise of authority	Personal attributes	Technical proficiency
Boys (N 57) age 11-13	15.8% (9)	40.4% (23)	43.8% (25)	26.3% (15)	35.3% (20)	38.4% (22)
Girls (N 59) age 11-13	15.2% (9)	52.6% (31)	32.2% (19)	23.6% (14)	42.4% (25)	34.0% (20)
Boys (N 120) age 14-15	15.0% (18)	44.2% (53)	40.8% (49)	18.8% (22)	45.4% (53)	35.8% (42)
Girls (N 106) age 14-15	16.3% (17)	59.0% (62)	24.7% (26)	17.0% (18)	63.2% (67)	19.8% (21)
Diff.	NS.	$p < 0.05$	$p < 0.01$	NS.	$p < 0.01$	$p < 0.01$

The exercise of authority by parents and other adults was by no means invariably mentioned with disapproval. In completing the sentences beginning: 'Mothers/Fathers/Grown-ups should . . .' reference to authority was made in approximately one-fifth of the statements. Younger boys made significantly more disapproving references than older boys, and older girls significantly more than older boys; but there was no difference between the younger boys and girls (Table 5).

Although the social distance scale and the choice-of-companions questionnaire showed the expected increase in independence of adults with age, there was no

Table 5. Mention of authority exercised by adults, mothers and fathers

	Approving (e.g. 'Grown-ups should not spoil their children', 'Mothers should learn their children manners', 'Fathers should make you do as you're told')	Disapproving ('Grown-ups should not boss us about', 'Mothers should not shout at us', 'Fathers should not hit us')
Girls		
14-15 yrs	24	35
Boys		
14-15 yrs	40	25
	$\chi^2 = 5.4$, d.f. = 1, $p < 0.05$	
Boys		
11-13 yrs	17	27
Boys		
14-15 yrs	40	25
	$\chi^2 = 5.5$, d.f. = 1, $p < 0.05$	
Girls		
11-13 yrs	12	13
Boys		
11-13 yrs	17	27
	Diff. n.s.	

corresponding increase in hostility and disapproval expressed in the sentences beginning: 'Grown-ups/Mothers/Fathers are ...'. One hundred and seventy-eight boys and 170 girls between the age of 11 and 15 (inclusive) completed these three sentences; all the sentences (and also the sentences beginning 'Boys (or girls) of my age are ...') were classified 'favourable', 'intermediate', or 'unfavourable'. Thus 'Mothers are wonderful' was classified favourable, 'Fathers are big-headed/always at the pub' unfavourable; sentences which simply stated facts ('Mothers are grown up') or a halfway attitude ('all right sometimes, not at others') were classified intermediate. Similarly with statements about their own age-sex group; statements such as: 'Boys of my age are bullies' were classified unfavourable, 'looking for jobs/still at school' intermediate, 'sensible/good pals' favourable.

Boys aged 15 made the lowest proportion of favourable statements (31 per cent) about fathers. (At age 14 66.6 per cent of their statements were favourable: C.R. 3.9, $p < 0.001$.) With the girls the lowest proportion of favourable statements about mothers (73.7 per cent) came a year earlier, at age 14. (At 15, 86.8 per cent of their statements were favourable: C.R. 5.1, $p < 0.001$.) On the other hand, there was no corresponding increase in either boys' or girls' favourable statements about the parent of the opposite sex (nor was there any transfer of approval to their own age-sex group) (Table 6).

There was no social-class difference at any age in the proportion of favourable attitudes to their elders. In Midland town the social range was not great; in the suburban school it was wide enough to make a comparison possible. 53.3 per cent of the statements made about adults by boys from white-collar and supervisory homes were favourable, 53.8 per cent of the statements made by boys from the families of manual workers; 80 per cent of the statements made by the former about mothers were favourable, 92.3 per cent of those made by the latter; 60 per cent of

Table 6. *Percentage of favourable statements about grown-ups, mothers, fathers, and peers*

	Grown-ups	Fathers	Mothers	Own Age-Sex
By boys				
11-13 yrs				
N 57	40.3	51.0	70.0	42.1
14 yrs				
N 63	49.2	66.6	77.7	49.2
15 yrs				
N 58	38.0	31.0	72.4	38.0
By girls				
11-13 yrs				
N 60	48.3	56.6	85.0	61.6
14 yrs				
N 57	44.0	49.1	73.7	35.1
15 yrs				
N 53	39.6	37.7	86.8	66.0

the statements made by the former about fathers were favourable, 77 per cent of the statements made by the latter. The differences between these percentages were not statistically significant. There was a similar measure of agreement between the statements made by girls from the homes of white-collar and of manual workers.

(d) *Adults' attitudes to adolescents*

The questionnaire which had been drawn up for adults was sent to 145 men and women in Midland town and to 125 in suburbanized village. Forty-seven (32.4 per cent) were returned from the former, 24 by women and 23 by men. Twenty-two of the women (91.6 per cent) were married, and all of the men; 15 of the women (62.5 per cent) and 17 of the men (73.9 per cent) were or had been parents of teenage children. The average age of the men was 53 years, of the women 50. Forty-three questionnaires were returned by the village (34.4 per cent). The 24 men, whose average age was 45 years, were all married and 46 per cent were or had been parents of teenage children; 18 of the 19 women who replied (average age 44 years) were married and 15 (79 per cent) had teenage children.

The sample drawn from Midland town was slightly biased towards the higher occupational grades, and more pronounced in its bias towards the higher age ranges. (Since the village is a civil parish within a rural district, it was not possible to obtain data from the Census relating to the age and occupations of the population.) Married women were assigned to the same social class as their husbands. Whereas 25.5 per cent of the sample were in Classes I and II, 51 per cent in Class III, and 23.5 per cent in Classes IV and V, the corresponding percentages for males over the age of 15 were 12.5 per cent, 54.8 per cent and 32.7 per cent ($\chi^2 = 7.6$, d.f. = 2, $p < 0.05$).

The sample was significantly older than the population of origin: 2.1 per cent were aged 20 to 29, 31.9 per cent 30 to 44, and 66 per cent over 45. The corresponding proportions for the adult population were 19.6 per cent, 33.6 per cent and 46.8 per cent ($\chi^2 = 11.04$, d.f. = 2, $p < 0.01$).

The analysis of the questionnaires showed no difference in type of response between the two areas when social class was held constant. The two samples are therefore combined, giving 45 people (50 per cent) in Classes I and II, 29 (32.2 per cent) in Class III, and 16 (17.8 per cent) in Classes IV and V.

The completed sentences beginning 'Teenage boys are . . .' and 'Teenage girls are . . .' were placed in three categories: (1) wholly or mainly critical and disapproving, (2) intermediate or neutral, and (3) wholly or mainly favourable and approving. The following are examples of statements in the first category: 'Teenage girls are conceited and selfish/excessively frivolous/far too impressionable/guilty of widespread immorality: it is horrifying how many girls have had intercourse in their teens, even as infants'. 'Teenage boys are arrogant and disrespectful/discourteous and pig-headed/sexually inclined and pleasure-seeking/given to theft, lying, idleness and disobedience—much more so than in my own youth'. The following statements were placed in the third category: 'Teenage girls are more intelligent than in former generations/as sweet as they were 50 years ago/attractive/more intelligent than boys'. 'Teenage boys are bright and clever/sounder mentally and physically than ever before/idealistic and full of energy/more sensible than girls/mostly well dressed and well behaved'. Intermediate statements were: ' . . . willing workers but foolish with money/all right provided they have proper guidance/very difficult to understand/confident but materially minded/pleasant but too advanced for their age'.

There was no tendency among the 90 men and women who returned the questionnaires for attitudes to teenagers to vary significantly with age, sex or family composition. No social-class differences were apparent in attitudes to teenage girls, but men and women in Social Classes I and II were significantly less favourable in their attitude to teenage boys than were men and women in Social Classes III-V. Only seven out of 48 men and women in Classes I and II (14.6 per cent) expressed a wholly or mainly favourable attitude to boys, but 16 out of 42 men and women in Classes III-V (38.1 per cent) did so. (C.R. 2.5, $p < 0.02 > 0.01$.)

The 540 sentences which had been completed were examined for dominant themes. Characteristics were most frequently ascribed to teenagers in four areas of behaviour: (1) rate of social-personal development: precocity, (2) tendency to be easily influenced and led, (3) appearance, (4) attitude to work. Teenage behaviour was most frequently prescribed in three areas: (1) helping in the home (most frequently for girls), (2) more training in housecraft (girls), (3) submission to firmer discipline (boys and girls) (Table 7).

No social-class differences were discernible in the distribution of these statements. Thus both men and women at all social levels made disapproving allegations of girls' 'sophistication', desire to appear older than their years and generally precocious development. Equally stern references were made at all social levels and ages to the teenager's need for more discipline: 'Teenage girls should be given far more discipline, should be in by 9 o'clock, should have less money, and a sound thrashing for screaming at fools like Elvis Presley' (31-year-old lathe-turner). 'Teenage boys should have far more attention paid to the formation of their

Table 7. *Ninety adults' perceptions of adolescents*

	Boys	Girls
Ascribed behaviour		
Precocious ('forward', 'too sophisticated' etc.)	Nil	34.4% (31)
Badly dressed ('slovenly', 'over-dressed' etc.)	2.2% (2)	13.3% (12)
Easily led and influenced ('sheep-like' etc.)	7.7% (7)	15.5% (14)
Lazy ('workshy', 'feckless' etc.)	33.3% (30)	10.0% (9)
Prescribed behaviour		
Should help in or about the house	3.3% (3)	12.2% (11)
Should have more discipline	23.3% (21)	15.5% (14)
Should have more training in housecraft	Nil	17.7% (16)

characters. Very firm handling of the large numbers of hooligans and gangsters roaming our country would stop their atrocious assaults on defenceless and elderly people' (80-year-old retired Brigadier).

The adults in this survey were far more hostile and critical in their attitudes to adolescents than were adolescents to adults (Table 8).

Table 8. *Adults' attitudes to adolescents and adolescents' attitudes to adults*

180 statements by 90 adults		
'Teenage boys are . . ., Teenage girls are . . .'		
Wholly or mainly favourable	Intermediate	Wholly or mainly critical
23.9% (43)	9.5% (17)	66.6% (120)
231 statements by boys and girls aged 14-15		
'Grown-ups are . . .'		
42.8% (99)	20.9% (48)	36.3% (84)

$$\chi^2 = 37.3, \text{d.f.} = 2, p < 0.001$$

DISCUSSION

Our prediction that between later childhood and early adolescence boys and girls would increasingly prefer their coevals to their parents as leisure-time companions was supported by the inquiry; but their movement towards independence of parents was less uniform than had been expected. In the case of both boys and girls a higher percentage chose parents rather than peers at age 11 than at age 10. No light is thrown by the study on the reasons for this: it is possible that the uncertainties which are associated with the onset of puberty, or even the disturbances associated with transfer from junior to secondary schools, are in some measure responsible.

In the light of previous research in this field, perhaps the most interesting results are those relating to changes in favourable and unfavourable attitudes towards parents. The stronger disapproval of adult authority expressed by pre-pubescent boys and by post-pubescent girls than by any other age-sex group is probably susceptible to a social-cultural explanation. The younger boys, turbulent even in their own eyes, and the older girls, 'forward' at least in the eyes of adults, probably

feel the constraints of adult authority more than other children, and have greater reason to refer to it with disfavour.

The peak period of hostility towards the parent of the same sex, 15 years for boys and 14 for girls, is broadly in line with Liccione's findings in America and confirms our original expectation. On the other hand, at no age did either boys or girls indicate the high degree of hostility to mothers that Liccione found among American adolescent girls. A far greater proportion of hostile and critical statements was made by both boys and girls of all ages about fathers than about mothers. Although 14-year-old girls made fewer favourable comments about mothers than older or younger girls, no less than 73.7 per cent of their comments were still favourable. Liccione argued that his findings could be adequately interpreted in the light of social-cultural theories as well as psychoanalytical frames of reference. The American mother probably pays the price for taking a dominant, disciplinary role within the family; English mothers are perhaps at an advantage in not enjoying a similar distinction.

This inquiry serves to underline the dangers of transposing American research on the family to English society. In the light of American work we supposed at the outset that boys and girls would sharply differentiate between the roles of father and mother. Pre-pubescent boys and girls appeared to make little distinction; neither did adolescent boys and girls, but by now their viewpoints had diverged. It is likely that both mothers and fathers behave similarly to their adolescent boys, and similarly again to their adolescent girls. The differences are between children of different sex, rather than between mother and father.

If there are dangers in transposing American work on the family to England, there are still greater dangers in transposing American work on the adolescent peer group. We began with the expectation that adolescents would show increasing hostility towards parents and increasing approval of their peers. There was little support in this inquiry for the often alleged 'solidarity' among adolescent males. Their approving statements about their peers were no more frequent than their approving statements about adults in general, less frequent than their approving statements about mothers, and up to the age of 15 less frequent than their approving statements about fathers. They were less frequent than girls' approving statements about their own age-sex group. It is possible that the very intensity of adolescent male peer-group association may be a cause of friction and hostility among them.

Our prediction that adults would more commonly express disapproval of adolescents than adolescents would express disapproval of adults was fully supported by the inquiry. Young adolescents showed themselves better disposed towards adults than adults were disposed towards them. The upper age limit of the adolescents who completed the questionnaires was, however, 15-16 years; whereas the adults were making statements about the whole range of 'teenagers'. It is possible that older adolescents are more critical of adults; although psychoanalytical theory would lead us to expect that the peak period of hostility, at any rate to parents, would be passed soon after puberty. The social circumstances of older adolescents, their greater social and economic independence and their increasing

involvement in the affairs of adult life, might also lead us to suppose that they would be less rather than more hostile towards their elders.

The picture which adults had of teenagers was widely different from the picture that adolescents had of themselves. The adults' picture was overwhelmingly negative, with scarcely any reference to teenagers' increasing social and technical competence. The Eppels found that the 135 responsible adults in positions of authority who replied to their inquiry were often unwilling to make general statements in reply to the questions they were asked; thus when they gave their views of the morals of young people today compared with the morals of adolescents in their own youth, 56 per cent 'withheld judgment'. But 29 per cent were generally disapproving in their remarks, and only 15 per cent appear to have been generally approving. This proportion of approving statements is not dissimilar in magnitude to the proportion in the present inquiry (24 per cent)—although it refers to only one aspect of adolescent behaviour, moral conduct. One positive quality which the responsible adults saw in young people today was frankness and intellectual honesty, which was mentioned with approval by 22 per cent of the sample.

Although in some respects the study reported above suggests important differences from American conditions, it is in line with those American investigations which have shown adolescents belittled by their elders, regarded as a separate, inferior, and even threatening population, exposed to contradictory expectations and demands from the general body of adults, and consigned, as Hollingshead (1949) has said, to 'an ill-defined no-man's land that lies between the protected dependency of childhood, where the parent is dominant, and the independent world of the adult, where the person is relatively free from parental controls'.

REFERENCES

ARIÈS, P. (1960). *L'Enfant et La Vie Familiale*. Paris: Plon.

BANTON, M. (1959). *White and Coloured*. London: Jonathan Cape.

BOGARDUS, E. S. (1947). The measurement of social distance. In *Readings in Social Psychology* (ed. T. M. Newcomb & E. L. Hartley). New York: Henry Holt.

BRONSON, C. W., KATTEN, E. S. & LIVSON, N. (1959). Patterns of authority and affection in two generations. *J. abnorm. soc. Psychol.* **58**, 143-152.

EPPEL, M. & E. M. (1962). Connotations of morality. The views of some adults on the standards and behaviour of adolescents. *B. J. Soc.* **13**, 243-263.

HALL, STANLEY (1904). *Adolescence*, vol. 2. New York: Appleton.

HOLLINGSHEAD, A. B. (1949). *Elmtown's Youth*. New York: John Wiley.

JERSILD, A. T. (1957). *The Psychology of Adolescence*. New York: Macmillan.

LICCIONE, J. V. (1955). The changing family relationships of adolescent girls. *J. abnorm. soc. Psychol.* **51**, 421-426.

LIPSET, S. M. & LOWENTHAL, L. (1961). *Culture and Social Character*. Glencoe, Ill.: The Free Press.

LUCAS, M. C. & HORROCKS, J. E. (1960). An experimental approach to the analysis of adolescent needs. *Child Developm.* **31**, 479-487.

MEAD, M. (1948). Social change and cultural surrogates. In *Personality in Nature, Society and Culture* (ed. C. Kluckhohn & H. A. Murray). New York: Knopf.

MORRIS, J. F. (1958). The development of adolescent value-judgments. *Brit. J. educ. Psychol.* **28**, 1-14.

MURPHY, G. (1947). *Personality*. New York: Harper.

NIMKOFF, M. F. (1942). The child's preference for mother or father. *Amer. sociol. Rev.* **7**

PARSONS, TALCOTT & BALES, R. F. (1956). *Family Socialization and Interaction Processes*. London: Routledge.

PITTS, J. (1960). The family and peer groups. In *An Introduction to the Family* (ed. N. W. Bell & E. F. Vogel). London: Routledge.

RIESMAN, D. (1950). *The Lonely Crowd*. New Haven: Yale University Press.

RILEY, M. W., RILEY, J. W. & MOORE, M. E. (1961). Adolescent values and the Riesman typology: an empirical analysis. In *Culture and Social Character* (ed. S. M. Lipset & L. Lowenthal). Glencoe, Ill.: The Free Press.

SHAW, H. (1954). Popular and unpopular children. *Educ. Rev.* 6, 209-220.

Manuscript received 3 November 1962

The Meaning and Measurement of Neuroticism and Anxiety: Supplement to a Review

BY RAYMOND B. CATTELL

Department of Psychology, University of Illinois

A good review commonly brings out clearly what new experimental or theoretical contributions the author has to make and then critically evaluates them. Unfortunately, Beech's recent and apparently maiden review effort on the above book dispenses with the first. Somewhere in his peer group he has heard a voice cry 'Havoc', and has so eagerly let loose the hounds of destruction that he has forgotten to describe the victim. On behalf of my fellow author, Dr Ivan Scheier, and myself, I should like to state soberly, and in the brief space allotted, the theoretical points and factual findings at issue.

First, we show that of the approximately twenty known personality factors in objective tests, now replicated ten times over in the general population, no fewer than six distinguish neurotics from normals at the $p < 0.01$ level of significance. In addition to the anxiety (U.I. 24) and regression (U.I. 23) factors which have been found before to do so, our study thus adds four of considerable theoretical importance, two of which have been found in nature-nurture studies to have appreciable hereditary determination. Surely the vital point for a reviewer to bring out here is the difference of formulation from the only other substantial work in this area—that of Hans Eysenck. As is well known, the latter presents a unifactor theory, calling U.I. 23 the 'neuroticism' factor, instead of the *regression* component, i.e. one influence among six as hypothesized here. This is the kind of clear conceptual difference, operationally investigatable, the bringing out of which can stimulate constructive research.

Secondly, these experimental findings are brought into a set of explanatory principles of neurosis formation, quite different from the psychoanalytical system, involving the Adjustment Process Analysis chart. To a reviewer competent both in clinical psychology and mathematical models this, as some reviewers have noticed, constitutes a remarkable challenge, offering an entirely new means of handling, by matrix algebra, the general problem of personality learning theory, of which therapy is a part. (Perhaps we should charitably assume that the pages of Chapters 11 to 13 got stuck together in this reviewer's copy.)

Defining personality learning as a multidimensional change in response to a multidimensional situation, and adding a third category of *integration learning* to *classical* and *operant conditioning* (so far constituting the stock in trade of reflexology), the new learning theory proceeds to matrix formulations. (In physics Heisenberg's matrix mechanics was momentarily received with similar blankness by classical physicists.) On p. 341 the total design of multiplying a matrix of *path-personality coefficients* by one of *path-frequency records*, to obtain a matrix of personality changes on source trait dimensions, is set out. Since personality factor batteries, such as the Objective-Analytic test battery, are now available to provide tolerably reliable scales for up to 18 of these dimensions, it is possible for an enterprising researcher to try out these equations wherever clinical records can reliably be quantified into a *path-frequency matrix*.

Possibly, to some practicing clinicians the above may seem a little too phantastically in the future, or too concerned with reformation of pure learning theory, to justify a reviewer's time. But the rest and third major novelty comes much closer to home. If an anxiety measurement is likely to be as common in clinics as is a glance at the clinical thermometer, it is important that the clinical psychologist be clear about what he is measuring. This book presents the first proof (assembled from the scattered specialized articles and two monographs of the Illinois laboratory) of the determinacy of the general anxiety factor and of the means of distinguishing it from factors with which it has hitherto been systematically confounded. The Taylor Scale, for example, by its derivation from items distinguishing neurotics from normals, mixes various other neurotic dimensions into anxiety. Eysenck's neuroticism scale, on the other hand, mixes too much anxiety into neuroticism. Of course, it may be

fairly argued, in relation to the latter, that since the clinically defined neurotic differs from the normal significantly on at least six factors on the 16 P.F. test (Table 6-1, p. 112, *op. cit.*), that the particular weights on these, e.g. on ego weakness, premsia, etc., are likely to vary with the culture. However, the weights given to these in building the brief IPAT Neuroticism Scale Questionnaire have some claim to stability, being derived from a wide sample, diagnosed independently in clinics a thousand miles, in several directions, from the Illinois laboratory. And the NSQ has a significantly and decidedly lower correlation with the IPAT Anxiety Scale (which is almost dead on the anxiety factor axis) than does the Eysenck Neuroticism Scale. This difference is scarcely to be accounted for by cultural sampling, for the finding which has astonished us (as sceptics of the psychiatric diagnosis process) is the remarkable absence of *any* significant differences on the 16 P.F. profile among neurotic groups so diagnosed in America, Australia, Britain and Canada.

The determinacy of the anxiety factor rests on two kinds of evidence sufficiently presented in this book. First, it can be fixed as a second-order factor among 16 P.F. factors, just as Thurstone fixed the general intelligence factor for the first time by means of his primary abilities. Secondly, it is fixed as a first order factor among *objective* tests, by a hyperplane defined in up to 18 dimensions (where three were the most previously offered). These are perhaps highly technical issues, concerned with the instability of inverses from factorizations employing only two or three dimensions, but vital conceptual issues in science hinge on complex technicalities, as, for instance, in the Michelson-Morley experiments. And, surely, the duty of the reviewer is to bring these into focus for the general reader. It is certain that the measurement confusion of anxiety with neuroticism and with certain other factors (notably with U.I. 20, Comention, by psychiatrists, see Table 5-5, p. 100) is going to vitiate and waste a great deal of current experimental work unless corrected.

If the confusion is as great as here indicated, in regard to anxiety as a trait, it is yet greater in regard to anxiety as a state. As a state, anxiety must be located by factoring change scores, in P-technique and in Incremental R-technique, and since there is dispute among mathematical statisticians about the sizes of standard errors in these methods the only dependable answer to doubts is replication of researches. Consequently, although the first P-technique studies on anxiety and other factors were published in 1947, it was not until 1961, in this book, that we felt justified in integrating firm conclusions. Among them are, (a) that the anxiety *state* factor is little different from the *trait* factor pattern, and (b) that anxiety state level measures have frequently been heavily contaminated with measures of activation level (P.U.I. 1, in the state index system) and effort stress (P.U.I. 4). One subject of our book was to show physiologists that the anxiety state could be validly measured, and another was to indicate a dozen or so of the chief physiological associations which need expert further investigation. In view of the current acute interest of physiologists and the general public in anxiety and coronary disease, the clear 'No' given here to any correlation of anxiety with cholesterol level, and the suggestion that the latter belongs in the distinct *effort stress factor* pattern, is perhaps worth noting. If explained by the reviewer, it could encourage some young researchers to follow this lead toward better definition, and separate measurement of, the effort stress factor, with resulting contributions in many directions.

The fourth and last of what may be the major novelties in this book (I hesitate to say contributions in view of the reviewer's high personal standards of how much research constitutes a contribution) is the carrying of personality factor analysis into second order and (in an aftermath) to third order resolutions. The data is based on about 10 hours of testing per individual, covering more than 80 diverse objective group and individual tests. The congruence of the independent 'blind' rotations is high and yields seven second order factors (p. 82) which clinicians may be pleased to find have a close resemblance, in three, to the famous trinity of the psychoanalysts, though the others lie in regions unknown to the elders of clinical tradition. Since much speculation has been to the effect that different factors analysts have been working at different order levels, this first full and replicated exploration of the pyramid to the third and last order should narrow the experimental issues. It may help meet certain justifiable criticisms that the factor analytic definers of personality structure should set their own house in order before attempting to explain clinical phenomena in their terms.

Any appraisal of the views in this book should also consider the findings since. (Two years elapsed between book and review.) In no major respect have the several articles since

published disagreed with the findings here described. Indeed, they have greatly strengthened and refined what was there put forward only tentatively, notably in a new study of higher order personality factors, and in the distinction of anxiety from neuroticism and from effort stress and other states.

When one of the greatest needs of our time is to transcend those differences in professional education which block the integration of clinical psychology, on the one hand, with the resources of psychometric personality structure analysis, on the other; when one or two books, at most, appear in this area in a year (as contrasted with two or three hundred of the Freud-to-Fromm variety) one might hope that a thorough airing would be given, in knowledgeable circles, to the strange ideas and methods involved. Instead, this book seems merely to have been exposed to a piece of juvenile vandalism.

Book Reviews

The Experimental Psychology of Beauty. By C. W. Valentine. London: Methuen, 1962. Pp. xi + 438. 43s.

Professor Valentine summarizes a selection of literature from 1880 to 1957 concerned primarily with a practical approach to the study of what is considered to be pleasing or beautiful. Examples include simple colours, colour combinations, pictures and *objects d'art*, poetry and music. The following topics illustrate the modes of approach considered. The comparative roles of 'elements' versus 'wholes'—The 'stimulus orientated' attitude of professional critics in contrast to the psychologists' emphasis on a number of different aspects of the *total* situation—The distinction between 'objective' and 'subjective' attitudes and between 'character' and 'associative' types of judgment—Personality types and preference—The scope of the conception of 'a general factor of taste'—Evidence from selected cross-cultural investigations.

The treatment is predominantly discursive and systematic theoretical examination is largely avoided, apart from a brief selection of comments by philosophers, art critics and other specially interested persons. Valentine makes his own personal orientation explicit from the beginning, which is in the diversity of individual responses to the same stimulus material. There are some rather surprising omissions from the references, discounting those which have appeared since the book was written and which illustrate the active development of interest in the general subject of 'Aesthetics' in recent years. Professor Valentine has performed a useful service in bringing together a wide range of literature concerning a practical approach to problems of aesthetics, but to refer to this as *The Experimental Psychology of Beauty* seems something of an overstatement. Certainly the book will be widely read, probably by a larger number of non-psychologists and psychologists. The detailed treatment of relevant statistical points is omitted, apart from occasional footnotes and a rather unsatisfactory appendix on this subject. Perhaps this book marks the end of one era and the beginning of a new one in so far as one feels that future methods of throwing significant light on these problems of aesthetics should develop along new lines rather than in terms of a repetition of the old.

ROLAND HARPER

Social Behaviour. Its Elementary Forms. By G. C. Homans. London: Routledge & Kegan Paul, 1961. Pp. 404. 30s.

In his previous work, *The Human Group*, Professor Homans was mainly concerned with establishing generalizations describing social behaviour, but in the present volume he boldly attempts to deduce these empirical generalizations from explanatory principles. These principles, a strange mixture of Skinner's operant conditioning and elementary economics, are formulated as five propositions concerning 'human exchanges'. Two characteristic examples of these are: 'The more often within a given period of time a man's activity rewards the activity of another, the more often the other will emit the activity,' and 'The more valuable to a man a unit of the activity another gives him, the more often he will emit activity rewarded by the activity of the other'.

The origins and simplicity of these principles may appear to make them more suitable for the explanation of the behaviour of pigeons in a market place than for dealing with the complexities of human social behaviour, but Professor Homans makes their application more plausible by severely restricting the type of social situation which he considers in this book. For he takes 'elementary' social behaviour to be 'an exchange of activity, tangible or intangible, and more or less rewarding or costly, between at least two persons', and only

considers this activity in so far as it is rewarded or punished by another person present at the time it occurs. Though the norms and institutions forming a background to such behaviour are admitted to be relevant, they are not in themselves Professor Homans' chief interest, and hence questions as to why a person was given power, or how a particular institution arose are ignored. However, he argues that his restricted 'elementary or sub-institutionalized' social behaviour is at the basis of all institutions, since its characteristics are universal while those of different institutions are not.

In successive chapters the author re-analyses, in terms of his five principles and their corollaries, the previously published results of experiments conducted by psychologists on influence, conformity, competition, esteem, and other aspects of social behaviour. As is to be expected, in these analyses the hunt for the 'reward' is at times strenuous, while the term 'activity' is made suitably flexible. Indeed, as he himself disarmingly admits, 'One difficulty in definition the pigeon's peck does not share with many human activities. Compared with them, it is pretty sharply discriminated both as a unit and as a type of activity. But even the peck is not an utterly standardized thing: the pigeon may peck hard or gently, and as for human activity, men may not only change the amount of it they put out but also change the activity itself by insensible gradations into one of another kind, without our being able to draw anything but an arbitrary line where one stops and the other begins.'

This account may give the impression that Professor Homans has been rather naive in his attempts at explanation. But this is not so, for when the pigeons are forgotten we are given brilliant insights into various facets of social phenomena. It is only a pity that such a distinguished sociologist should have chosen, out of all the current psychological theories and concepts, a simplified version of Skinner's operant conditioning in order to try and explain social behaviour.

A. R. JONCKHEERE

Social Judgment. Assimilation and Contrast Effects in Communication and Attitude Change. By Muzafer Sherif and Carl I. Hovland. New Haven and London: Yale Univ. Press, 1961. Pp. xii+218. \$6.00.

This is the fourth volume of Yale Studies in Attitude and Communication. Carl Hovland was the general editor of the series, and it is sad to record that this is a posthumous book. He died in 1961 after a severe and protracted illness. He was already very ill in 1959; but I remember him contributing actively and incisively to a discussion at a colloquium concerned with problems very close to those with which this volume is dealing.

The question which seemed to preoccupy Hovland in 1959 is still of central relevance to a consideration of the volume on *Social Judgment*. The discussion at the time followed after a paper which presented a scheme of empirical generalizations about the judgmental aspects of social perception. Hovland's contributions to the discussion can be summarized as follows: It's all very well to attempt a series of general statements relating shifts and biases in the judgment of physical objects to similar shifts and biases in the judgment of social objects; but why should these similarities occur? *What can we say about the underlying judgment processes?*

This is the central issue of the book. The title of Part I is 'Leads derived from studies of judgment'. In it the authors summarize a number of experiments, some of them their own, concerned with the effects of 'subjective' variables on judgment. This summary can hardly be bettered. The displacements in judgment determined by various types of 'anchorages', internal and external, are shown to be reducible to a small number of general laws allowing for an efficient prediction of a great variety of systematic constant errors.

The second part of the book is an attempt to apply these general laws to problems of communication and attitude change. With the help of illustrative experiments, Sherif and Hovland analyse the relation between an individual's position on a given social issue and the effects on this position of various communications pertaining to that issue. This analysis is made in terms of concepts derived in large measure from the first part of the book: choice of anchors in unstructured judgment situations; effects of these anchorages on categorization of statements expressing different degrees of agreement or disagreement with the subject's

own position; effects of various distances on the scale between the subject's position and communications attempting to shift it in one direction or another.

The predictions about attitude change which follow from this analysis are something of a mixed bag: many will be of direct interest to any hidden persuader, some are near-platitudinous ('We infer that few subjects with extreme positions can tolerate alternative views on the issue, at least when the topic is highly ego-involving'). But the importance of these occasional trivialities should not be over-emphasized: more important is the emergence in the book of a system of generalizations applying to a fairly wide range of phenomena in the field of attitude change.

What of Hovland's question about the underlying judgment processes? We still seem to be far from an answer: the fact that anchorages seem to determine similar effects in psycho-physical and psycho-social judgments does not tell us *how* these anchors work. Perhaps in order to find an answer we should put the cart before the horse; it may prove rewarding to try and derive laws of shifts in judgment from the analysis of the process of categorizing rather than the other way round. In such an attempt, displacement of judgments on physical scales need not be considered as a basis from which all other related phenomena must be inferred. When such displacements are more closely related to the categorizing activities of the subject than to the physical properties of the stimuli, they need not be considered as more 'primary' or 'basic' than similar displacements observed in the study of attitudes, or of perception of people, or of stereotyping. This view is perhaps not very far removed from the implications of some of the statements made by Sherif and Hovland in the section of the book devoted to 'Suggestions for further research'.

A pious and theoretically irrelevant hope to conclude: that the clockwork shifts in attitude predicted and described by the authors may prove less convincing in experiments using rational persuasion addressed to groups of rational subjects.

HENRI TAJFEL

The Acquaintance Process. By Theodore M. Newcomb. New York: Holt, Rinehart & Winston, 1961. Pp. 303. 6os.

In earlier work, particularly the well-known Bennington College study, Newcomb has demonstrated that increasing acquaintance leads to attitude change. In the present study, the emphasis is different, for attitudes remain fairly constant over time: it is shown therefore that an effect of similarity of attitudes held by interacting people is increasing attraction between them. The study is, however, a very detailed one, conducted on partly manipulated, partly naturalistic subject groups. In each of 2 years, 17 students unknown to each other were assembled to live together for a year in a separate house while in other respects pursuing their University studies in a usual fashion. They received certain advantages in return for regular sessions for completing attitude schedules, sociometric questionnaires and personality inventories. A great deal of data, permitting the testing of a number of hypotheses, was the outcome.

The hypotheses are formulated within a general theory of balance systems, having a family resemblance to several other current theories in social psychology, notably those of Heider and Festinger. The elemental system consists of two persons and an object (or person) about which each person has an orientation: imbalance or 'strain' is present when there is a discrepancy between one person's orientation and his perception of the other person's orientation towards the same object. The two main hypotheses are that for a single person, such systems will tend to remain in balance at all times; and that in systems of more than two persons, there will be similar pressures towards balance but increasing with length of acquaintance.

The main hypotheses are largely established, although there are certain discrepancies between the two subject groups. Subsequent sections of the book fill in some of the intervening variables, such as relevant individual differences, and test predictions about structural changes consequent upon increasing acquaintance.

In the main, then, this is a carefully programmed study, with the hypotheses and the experimental operations clearly stated, and with several appendices setting out the *rationale*

of the various statistical procedures adopted. It is not entirely without serendipitous findings, however, particularly in relation to differences among broad personality types. Of course, such a manoeuvred social setting permits a high degree of uncontaminated hypothesis-testing; but for wider generalization, it may prove necessary to extend the basic theory to take account of other alternatives open to the persons involved; i.e. something akin to the 'comparison level' notions of Thibaut and Kelley.

The book is far from lavish in production: rarely does one find these days so little space for *marginalia*. The price, therefore, seems unduly high.

THELMA VENESS

The Short-Term Prisoner. By R. G. Andry. London: Stevens, 1963. Pp. xxii + 162. 28s.

Dr Andry, whose first book (*Delinquency and Parental Pathology*, Methuen, 1960, 21s.) drew attention to the neglected role of the father in the genesis of juvenile delinquency, is now concerned with adult offenders. The great bulk of our prison population consists of males in their twenties who are 'inside' for the time and whose sentences are of less than six months. The Prison Commissioners claim that 80 per cent of these do not return to prison for subsequent offences. Dr Andry's main objective was to discover what distinguishes those who do return.

By means of a factor-analysis he reduced 22 variables to four factors which appear to be associated with 'recidivism', and between them account for 45 per cent of the variance. They are 'emotional immaturity', 'neuroticism', 'extra-punitiveness' and 'juvenile crime experience'. The three psychological variables are based on the interviewer's assessment; the penological variable on the prisoner's admissions. Dr Andry goes on to argue, from what is known or assumed about the psychological variables, that the only type of offender who is likely to benefit from a prison sentence is the 'mature, non-neurotic, extra-punitive' man, and he needs a rather longer sentence. As for the rest:

'During the course of interviewing the prisoners (and following up their lives on release, so far as it was practicable) the writer sensed that, for the majority of short-term offenders what is really wanted is a certain degree of reforming punishment, combined with therapy, coupled with an opportunity to change aspects of the personality through satisfying work—this being especially true of the immature and neurotic offender. . . .'

For them he suggests a 'Reconstruction Centre' staffed by the Probation Service.

Although Dr Andry is modest about what he called a 'small-scale exploratory study' these conclusions will undoubtedly be taken at their face value by most of his readers, and it is therefore important to examine their scientific basis. His sample was taken from the receptions at a London local prison, and amounted to 114 men, of whom 18 had 'returned to jail' by the time he decided to close his files, and were thus classified as 'recidivists'.

Clearly, the criterion of 'recidivism' is of crucial importance. What one would like to be able to count is the number of men who committed further breaches of the criminal law (or further indictable breaches, if we agree to exclude such trivial misdemeanours as keeping a dog without a licence). But we must be realistic and accept that some of any sample of offenders will commit further offences without detection. This being so, however, every effort must be made to eliminate all other avoidable kinds of bias. The follow-up period for each offender must allow him roughly the same period 'at risk' as the next man; and it must be long enough to catch the man who manages to avoid committing (or being detected in committing) offences for a short period, but relapses later. (W. Hammond's 11-year follow-up of 3,143 Scottish first offenders (to which Dr Andry does not refer, although it is the best published study of reconvictions) showed that of those who were eventually reconvicted only a third were reconvicted in the 12 months after their fine, imprisonment, etc.) Few investigators nowadays are content with a follow-up of less than 42 months. But on p. 2 Dr Andry reveals that his follow-up period ranged from 10½ months to 'less than 3 months', and he was content to regard a man who had not returned to prison as a non-recidivist, although it is by no means unheard of for a second indictable offence to be dealt with by means of probation.

Dr Andry discusses the other limitations of his study so frankly that it seems possible that he did not appreciate that a 2-year follow-up of reconvictions in the case of each offender would in all probability have added another 18 men to those whom he classifies as recidivists. Thus he has performed a factor analysis on 114 individuals, of whom 18 are correctly classified as recidivists, and an unidentified but roughly equal number are incorrectly classified as non-recidivists.

These defects do not, of course, destroy the interest of his proposal for a Reconstruction Centre, especially since he is the first to point out that the proposal is not really a necessary result of his findings. The idea should be discussed on its own merits, as it no doubt will be in penological journals. It should be realized, however, that if the policy of such a centre is to be based on a typology of recidivism, this typology has to be firmly established.

NIGEL WALKER

Roots of Behaviour. Edited by E. L. Bliss. New York: Hoeber, 1962. Pp. 340. 6 gns.

This is a symposium on many biological aspects of behaviour which has been contributed to by 31 authors. Although the papers were originally given as lectures at a meeting, it is a welcome relief not to have any details of discussion published.

The papers themselves fall into four categories concerning respectively Genetics, Instinctive Behaviour, the Effect of Early Experience on Behaviour, and Social Behaviour leading up to human activities. With so many authors concerned it is inevitable that the standard of different contributions should vary considerably. It is particularly noticeable, for example, that while some authors have been mainly concerned with general reviews of a particular topic, others have given detailed quantitative accounts of their own research. This is unfortunate since there is a real need for general reviews in the fields mentioned above and it seems to the reviewer that symposia of a general sort should not be the repository for detailed accounts of experimental results presented in a review paper, particularly when the original experiments have not apparently been reported elsewhere.

The four sections into which the book falls reflect fairly accurately the activity of current work in the field of comparative psychology, though they do not extend to those more biological aspects such as evolution or ecological adaptation. The importance of genetics in the study of behaviour is becoming increasingly obvious but is still at a very descriptive level of analysis. It is clear that the use of simple organisms of well established genetic structure, such as *Drosophila*, is more likely to throw light on the mechanisms of inheritance of behaviour, than studies on vertebrates.

The section on instinctive behaviour extends from mechanisms involved in reproductive behaviour, by Young and Lehrman, an account by Harlow and more recent work on the development of affection in monkeys, to studies of the development of suckling behaviour in kittens by Schneirla and his colleagues. It is clear from this last paper and from the section on early experience that the old distinction between innate and learned has become most infertile when studying the detailed development of behaviour. The important feature of the study of development seems to be in the interaction between the environment and the organism, without having preconceptions about the nature of learned or of genetic determination. It is also clear that the concept of imprinting is rapidly widening, since a variety of critical periods for a large number of different sorts of behaviour have been found, during which the organism is particularly sensitive to environmental effects. The final section on Social Behaviour is rather patchy: the two papers by Carpenter and Altmann reflect the present active state of studies of primate behaviour under natural conditions. The reviewer can make little comment on the paper called 'The Behavioural Sink'.

The book is well and even luxuriously well produced, which perhaps accounts for its high price. In some cases it will serve as a good introduction to many aspects of behaviour, while in others the papers are of sufficiently high standard for an Honours course in Psychology. It is certainly a useful book for teaching at different levels, provided one selects one's references according to the reader, and also a good source of reference to current work in comparative psychology.

D. M. VOWLES

Speech and Thought in Severe Subnormality. By Neil O'Connor and Beate Hermelin. Oxford: Pergamon Press, 1963. Pp. xi + 122. 25s.

Any contribution to our meagre knowledge of the large population of mental defectives is always welcomed, but a book of this calibre and approach is especially refreshing. This is a tightly written report of research into the character of the perceptual, cognitive and linguistic processes of imbeciles. It is described in the foreword by A. R. Luria as a work distinguished by the breadth of its scope and the careful systematic nature of the experiments reported. We are hardly disappointed.

The authors present first an experimental analysis of visual perception and discrimination in imbeciles. The evidence offered suggests that elementary perception is not as impaired as the imbecile's ability to focus on the relevant features of a visual display. These findings are followed by a broad study of the subnormals' linguistic abilities, and the effects of verbalization and verbal coding on their performance in discrimination experiments.

Characteristic of the imbeciles' behaviour in discrimination tasks is an apparent disinclination, as much as disability, to verbalize concepts. When forced to verbalize, however, their performance was marked both by improvement and increased stability of the reaction, clearly supporting the Soviet view of the directional function of language.

Continuing their exploration into the ways in which language might be utilized in controlling imbecile behaviour, the authors outline a series of experiments in which verbal coding was made an inherent part of a discrimination task. A higher level of discrimination was attained by the imbeciles when stimuli and responses were presented in 'opposite' sense modalities, where a verbal formulation and translation was required for the response, than in either a unimodal or mixed modality S-R relationship where no such verbal coding was necessary.

The nature of memory in imbeciles was examined with special reference to the effects of newly learned material on short and long term recall of previously learned items. Some encouraging observations concerning the imbecile's ability to benefit from incidental learning and to develop learning sets and transference were offered in this connection.

The final problem area to be treated was represented by a series of exploratory experiments aimed at a physiological and psychological comparison between mongol and non-mongoloid imbeciles of the same I.Q., and normal subjects, in such aspects as skin-conductance, EEG alpha wave blocking, eye-movements and stereognostic recognition. Differences between the groups on certain of these measures were shown, and tentative hypotheses offered to explain the findings.

Throughout the book, imbecile performance is compared with that of normal children of equivalent mental age.

O'Connor and Hermelin have attempted to demonstrate that deficits in certain areas, such as in acquisition and coding, are both more marked and important to the imbecile's efficiency than those in elementary perception, retention and transfer.

Moreover, they have offered an independent support of Luria's emphasis on the vital role of language as a 'second signalling system' modifying and controlling behaviour. The theoretical and educational implications of the imbeciles' use of verbal coding in discrimination problem-solving is evident throughout.

This work is characterized by a clear and logical presentation of the empirical data and theoretical interpretations, together with an almost religious regard for the relevant literature in the areas examined. It offers both a direction and a solid foundation for future investigations into the cognitive processes of the severely subnormal.

ARTHUR WINGFIELD

Theory of Collective Behaviour. By Neil J. Smelser. London: Routledge & Kegan Paul, 1962. Pp. 387 + Index and bibliography. 45s.

This most impressive book is a major contribution to the theoretical explanation of collective behaviour. Professor Smelser sets out to provide a morphology of types of collective behaviour, and to discuss the specific determinants of each type. He employs a

conceptual framework of some complexity, of the kind with which readers of Talcott Parsons have now become familiar, in which he seeks to distinguish seven levels at which values operate—from the most general societal level to the most specific case. Similar levels are described for the normative structure; for the mobilization of individual motivation; and for the situational facilities of social action. These are the components of social action. Thus at the most specific of the seven levels—using the example of economic behaviour—we have the values of work and personal responsibility, the norms of scheduled work programmes, the motivation of the particular role within the work organization, and the provision of specific facilities for its performance. At the most general level we have societal values, general conformity, socialized motivation of the individual, and prevailing ideas about causality. Our comprehension of the author's categories is facilitated by a series of tables and examples. Even so, it is not altogether clear just what assumptions are made about the logical relationships of his categories at their various levels.

Collective behaviour is thus categorized according to the point of this system at which it deviates from the formal pattern of social action. It is assumed to be a product of strain felt by the actors—no matter whether the strain is called anomie, disorganization, deprivation or role-conflict, or anything else. Collective behaviour is a response which seeks to reconstitute one or more components of social action in order to reduce or eliminate the strain which the actors experience. Thus the phenomena under review include not only panic, as a direct consequence of hysteria, but also the craze (wish fulfilment); the hostile outburst, which is directed against individuals or groups who are 'held responsible' for strain; the norm-oriented movement, which might be for social reform or an alteration of the law; and the value-oriented movement, which might be a religious or revolutionary movement seeking to transform the basic values of society.

He seeks by use of what he calls 'the logic of value-added' items to show how each type of collective behaviour is specifically determined. This approach is not a simple chronological reconstruction of events, but essentially an attempt to show how each stage of causal circumstance is a necessary, but not a sufficient, condition in determining the next; each stage narrowing the range of possibility of the ultimate outcome. Thus, given the experience of strain, collective behaviour is the attempt to reconstitute social action at some more general level so that the point at which strain is felt can be restructured and strain eliminated. Smelser does not suggest that there is a direct causal connection between the point at which strain is experienced and the type of collective behaviour, but that this relationship is mediated by intervening variables which he examines in terms of his value added analysis. One such variable is the generalized belief about the source of strain—such as rumour or ideology. Such beliefs typically 'short-circuit' the connecting links of the system of social action; for instance, the generalized belief that 'if only we all led purer lives' certain calamities would not occur, might lead to a value-oriented form of collective behaviour to avoid the (re-) occurrence of natural disasters.

Inevitably, in so ambitious a theoretical scheme, there are lacunae. For instance, his categories tend to exclude certain types of movement by definition. Such would be movements with general aims which, nonetheless, accept the broad value-structure of the society—conventional (as distinct from revolutionary) political parties, established churches and denominations (as distinct from sects), leagues and voluntary associations. Once a movement becomes institutionalized, or seeks to amend the values or norms of the society by institutionalized procedures, it passes beyond Professor Smelser's concern. Undoubtedly there are good theoretical grounds for this distinction, though at times it appears arbitrary. Obviously Professor Smelser has sought to avoid characterizing collective behaviour as dysfunctional, with all that that implies. And yet, in some measure, his frame of reference—and the use of strain as the crucial determinant—suggests that collective behaviour is necessarily deviant behaviour, since it seeks to alter the components of social action. At the level of scapegoating or lynching (the hostile outburst) or a hysterical demonstration (panic) collective behaviour is easily characterized as deviant, irrational and emotional, and so may the response 'there should be a law against it' (norm-oriented movement) but this is more likely to depend on the sympathies of the commentator. It is only fair to say that Professor Smelser avoids all value-judgments, but in treating lynchings and, say, incipient religious movements such as early Christianity or revolutionary parties as collective behaviour, it is unlikely that he will escape the censure of those who demand judgment of these phenomena in terms of the evaluation of the ideas which stimulate them. The reply, of course, is that sociology seeks

merely to explain in causal terms, that this model seeks to analyse similar types of social phenomena, and that it is no part of its business to distinguish in evaluative terms between the justifiability of spontaneous outbursts and long-contemplated revolutionary campaigns.

Professor Smelser provides a model of social action which necessarily takes 'as given' the value-structure and norm-structure of a social system. Those who seek to amend it proceed from some felt strain in the system. Thus, he does not discuss 'conflict' between those defending, and those attacking, the *status quo*. But as long as the model is recognized as such, and as long as Professor Smelser does not suggest that this is the true model for social analysis (and he does not) I do not think that his analysis can be regarded as prejudiced in favour of the *status quo*. Had he confined himself to lynchings and witch-hunts he would not excite such criticism. But applying his model to protest movements, some of which have strong ideologies about their own rationality (and some of which share the orientations and values which gave rise to sociology itself), he will certainly call forth hostility in some quarters. To this reviewer his position seems defensible, and his contribution to sociology one of real importance.

B. R. WILSON

A Treatise on the Medical Jurisprudence of Insanity. By Isaac Ray. Edited by Winfred Overholser. Cambridge, Mass. and London: The John Harvard Library; Harvard University Press and Oxford University Press, 1962, Pp. xvii-376. 54s.

The reissue of Ray's *Treatise*, first published in 1838, is a timely reminder that many of its problems remain unsettled. Ray was concerned with the criteria of criminal responsibility. 'Few . . . are aware how far the condition of the law relating to insanity is behind the present state of our knowledge concerning that disease.' This gap between the legal concept of insanity and the medical concept of mental disease is as wide today as in Ray's time.

Ray proposed to abandon the legally defined and intellectually oriented tests of 'right and wrong' (e.g. the 'McNaughton Rules' discussed in the 5th edition of the *Treatise*). He recommended that the jury advised by the testimony of experts should determine whether the unlawful act was a product of mental disease or defect. This implies the controversial proposition that the causal connection is a question of fact to be established by experts. Ray's opinion has influenced the 'New Hampshire Rule' (*State v. Pike*, 49 N.H. 399) providing an alternative to the English practice. The rule was reaffirmed in 1954 by the U.S.A. Court of Appeals for the District of Columbia quoting Ray as one of its authorities.

Ray agreed with Pinel that 'the affective as well as intellectual faculties are subject to derangement'. This led to the acceptance of 'moral mania' as a legitimate instance of insanity, including conditions now classified as psychopathic personality and also the irresistible impulse (admissible as a defence in Michigan, Georgia, Louisiana and some other States).

Other parts of the *Treatise* still of importance are ch. xiv on 'Lucid Intervals' and Appendix II (ch. xxix of the 5th edition) on 'Duties of Medical Witnesses'.

Much of Ray's work survived the test of time less well. Here belong his system of classification and his enthusiasm for phrenology 'the only metaphysical system of modern times that professes to be founded on observation of nature . . . '.

There were five editions of Ray's book in the U.S.A. and one in England. The present text follows the first edition with the most important changes included in appendices. It was carefully edited by W. Overholser who also wrote the introduction. A book worth having.

BERNARD MANDEL

A Model of the Mind. By Gerald S. Blum. New York and London: Wiley, 1961. Pp. xi+229. 53s.

The present reviewer must, he thinks, have been unfortunate in his recent diet of American books. It has sometimes seemed as if authors were competing for the distinction of wrapping up their meaning most successfully. Much of the writing is such that terms like

'abominable' are scarcely strong enough to describe it, and expressions like 'gibberish' hardly too strong. And yet there is no need for this state of affairs, as many well-written volumes testify. Professor Blum is not one of the worst offenders, but the present book is not one of these well-written volumes either. It is difficult to read and to follow, and of necessity the case which Professor Blum wishes to make is marred by his language and manner of presentation. This reviewer, after having read the book repeatedly at intervals, is still not sure just what Professor Blum really means. His account may therefore be doing Professor Blum considerable injustice; but this, the reviewer would plead, is at least partly Professor Blum's own fault.

The book is in three parts. The first part contains a short exposition of the theory; the second part reports a number of experiments using hypnotism, and relates these to the first part; the third part briefly interprets various common psychological phenomena in Professor Blum's language. Professor Blum is convinced of the 'ultimate feasibility of an all-embracing detailed theory of the mind's operation', in the direction of which he considers his theory to be a step. His model of the mind is essentially a pseudo-electronic model providing a sort of mental circuit-diagram. Three 'subsystems', the cognitive, the affective and the motoric (*sic*) make up the 'mental system'. The cognitive subsystem is the master subsystem which 'activates' the others. Incoming signals are processed by 'components'. A process of 'facilitation' or strengthening of 'connections' in the system enables subsequent signals to pass more readily. Separate signals are 'integrated' into composite signals. A process of 'hyperfacilitation' amplifies the strongest composite signals at any moment. 'Reverberation' represents positive feedback. The results of the cognitive processing affect the affective and motoric subsystems and lead to 'effector output' which, via 'response feedback', affects the 'sensory input to the cognitive subsystem'.

It is hoped that this will give the reader some idea of the kind of theory with which Professor Blum is concerned. Without going into details, the present reviewer would simply like to comment that, if a comprehensive theory of 'the mind' is attainable, he does not feel that it will be this kind of theory. He was more interested, as he suspects many readers will be, in the descriptions of the various procedures which were carried out under hypnosis, although finding the interpretations confusing. Perhaps he just has not read the book often enough to be able to decipher all its contents sufficiently clearly.

D. GRAHAM

Readings on the Exceptional Child: Research and Theory. Edited by Philip E. Trapp & Philip Himmelstein, London: Methuen, 1962. Pp. xii + 674. 55s.

This is a collection of papers published in recent years, and of some original articles by American university teachers, directors of clinics and other mental health workers, many of whom work both in the academic and applied fields. It is designed to acquaint students with some of the problems, theoretical approaches, and types of experimental design which characterize contemporary work. Earlier papers which have stimulated research are followed by more recent and controversial findings, and each section includes review articles. There are also some papers on conceptual and classificatory problems.

The largest section is devoted to mental deficiency, weighted somewhat on the Lewin-Kounin concept of rigidity with recent reinterpretations in terms of motivation. Work on visual discrimination learning, learning sets and reasoning in the mentally defective is represented in several papers. In comparison, there are relatively few articles on the gifted, blind, deaf, and speech-handicapped. The usefulness of handicapped groups for basic research is demonstrated in a paper on space perception. Reviews of experiments relating speech difficulties and personality variable provide correctives for overhasty generalizations. In the chapter on brain-damage and physical handicaps the notions of forced responsiveness to external stimuli and concrete behaviour of the brain-damaged are examined in studies that point up earlier faulty methodology and the heterogeneity of both the organic and familial groups. A number of papers deal with the speech development, adjustment and motivation of these children, but there is a dearth of studies on emotional disturbances without organic or sensory complications. Three of the six articles on exceptional emotional

processes deal with childhood schizophrenia, and there is only one on psychogenic disturbances—a clinical investigation of school phobia. Although this is a fair reflection of the proportion of experimental work appearing in each area, it is not clear why, in a book of this kind, there is no mention of the studies on anxiety and its relation to task difficulty and test efficiency. There are no examples of work on aggression or delinquency, and all etiological and psychometric studies are excluded on grounds of lack of space.

A book of readings cannot be exhaustive, and it is probably inevitable that certain problems should be disproportionately weighted, and the quality somewhat uneven. A more serious drawback is the lack of references to work in other countries. These criticisms, however, only emphasize that there is a need for books of this kind which collate recent work in a form easily available to students. The method of exposition, the emphasis on well-controlled experimental studies, the absence of theoretical bias, and the stress on a link between fundamental and applied work are to be welcomed.

SUSANNA MILLAR

Human Development: From Birth through Adolescence. By Ira J. Gordon. New York: Harper, 1962. Pp. 400. 46s.

This account of the development of a new-born infant into an adult attempts to combine an internal and an external approach. The author's standpoint is that recent research indicates the need for a rapprochement between the psychology of the individual's private world and general behaviour theory. He recognizes that the hypotheses generated by the former are difficult to test but looks to the latter for the necessary experimental rigour. In seeking the best of both worlds, he claims to have achieved a unifying theoretical position.

The book is mainly a review of the relevant American literature from 1950 onwards. The British literature is barely considered, while even Piaget attracts no more than half a dozen references in the text. This bias might be more acceptable if the author had attempted a more searching appraisal of the work which interests him; as it is he is apt to quote without criticizing, and there are large sections (including the summaries at the end of each chapter) which contain no information. Moreover, as in so many American texts of this kind the use of hyphenated jargon can have a paralysing effect; for example 'Since one of the causes of peer-group formation is the independence-from-parents motive of the members, peer-parent transactions are often struggles along the dominant-submissive axis' (p. 177). If the author means by this merely that children use their relationship with peers as a means of growing away from their parents, who may resent it, why not say so? The trouble is that one cannot always be sure what he does mean.

However admirable its intentions, this book appears to the reviewer to fall between two stools. It is too long to serve as an introduction to developmental psychology, while it is not sufficiently comprehensive to meet the needs of teachers and research workers. The author's sympathetic understanding of young people in their struggles towards self-fulfilment shines through every page, and the section on Preadolescence is especially commendable. For the reasons mentioned, however, he seems unlikely to command a wide circulation in this country.

MICHAEL HUMPHREY

Toward a Theory of Being. By Abraham H. Maslow. New York and London: Van Nostrand, 1962. Pp. 214. 15s.

For a decade now, books have been appearing with terms such as 'Altruistic love', 'Becoming' and 'Being' in their titles. It is difficult to learn a new language in psychology; even more when it resembles an old language, and one that has been discarded for scientific communication. And yet there must be many people who have experienced Maslow's disease with the 'reaction' psychologies and with an extreme plasticity view of human nature; people who would welcome a radical departure from these. They may find, however, that Maslow's book induces at best a love-hate reaction.

Maslow describes this book as a continuation of his *Motivation and Personality* and his general position as representing a 'Third Force' in psychology—an alternative to behaviourism and psycho-analysis. Along with many other writers, at least from Jung onwards, he stresses the self-actualizing processes maintaining that in the maturation of normal people, once the 'deficiency needs' are met, then other needs take over-needs which contribute to true happiness and fulfilment. These are deemed to be as much part of man's inner nature as his survival needs, and dependent on the survival needs only in the sense that their flowering is contingent on adequate provision for the latter. In treating the deficiency needs Maslow makes insufficient distinction between transitory states such as hunger and anger, and deficiencies of the kind Horney describes as 'neurotic needs'. Occasionally, he remembers that society could play a part at least in determining whether the survival needs are left screaming, but there seems little room for culture or social pressures in shaping the self-actualizing needs. These come entirely from within. This, then, is a thoroughly rosy view of human potentiality: and yet he has to admit that only very few people can be described as truly 'self-actualizing'.

Amplification and illustration of the main points comes from reports made by normal but articulate (and possibly privileged) subjects about their 'peak experiences', of mysticism, creativeness and love. The interesting suggestion is made (yet another attempt to give a scientific account of values) that from the study of people capable of such experiences—the peak of human development—and the life choices they make, could emerge soundly based recommendations about the good life. There would seem to be a need for a panel of highly selected specimens—a 'Sana' equivalent to 'Mensa' to provide the research material.

Maslow writes as a man very excited by his ideas, but a result is that what emerges often reads like notes for a book rather than the book itself. But what one misses most is reference to work which represents the hard grind of studying growth processes. Harlow just gets into the text but not into the bibliography; Piaget has no place at all. The bibliography of 187 items is explicitly devoted largely to writers deemed by Maslow to be exemplifiers of the 'Third Force'.

THELMA VENESS

Bodily Sensations. By D. M. Armstrong. London, Routledge & Kegan Paul, 1962. Pp. vii + 132.

This small book is a very clear and able piece of philosophical analysis. Armstrong sets out 'to give an account of the nature of bodily sensations'; and the book is an analysis of the concepts involved in this field.

Armstrong begins with a standard distinction between two classes of bodily sensations. With a sensation of warmth or pressure, we can distinguish the warmth or pressure from the sensation of warmth or pressure. But with a sensation of pain or an itch, we cannot distinguish the pain or itch from the sensation of pain or itch. The former class he calls transitive bodily sensations, the latter intransitive. He employs another well known distinction in epistemology—that between mediate and immediate perception—and uses this to introduce the notions of a *tactual* and a *bodily* sense impression. He concludes that 'all immediate *tactual* perception involves perception of a relation between our body and something in contact with it'. All the properties of objects immediately perceived by touch reduce to spatial properties, except for heat and cold. To talk of transitive bodily sensations is to speak of bodily sense impressions, or, in some cases, *tactual* ones. The intransitive bodily sensations, however, present a much harder task, Armstrong says. He reviews five different views about them. To have a bodily sensation, e.g. pain, (1) is to perceive a quality of the body; (2) is to apprehend a unique non-physical item located in the body; (3) is to apprehend an item that is not located in the body; (4) is just to take up a certain attitude towards a portion of the body and not to apprehend any item at all; (5) is to have a bodily sense impression. Armstrong comes down eventually in favour of the last alternative; and attempts carefully to meet the objections to it. On this view, to have a pain in a certain place is to have a bodily impression, namely that of feeling 'a disturbance of our normal bodily state at that place; together with an immediate and interested dislike of that feeling; and a concern for the place where the disturbance feels to be'. The upshot, therefore, is that

bodily sensations in general (transitive or intransitive) are a species of sense impression; and as they involve impressions of thermal and spatial properties alone, we do not need to extend any further the list of 'secondary qualities' in order to deal with bodily sensations.

I think Armstrong has given us a valuable piece of elucidation. He has obviously written the book with his eye primarily on a philosophical audience, not a psychological one. But if any psychologist is puzzled and feels confused by his own ordinary language, or by his semi-technical or technical talk, about touch, phantom limbs, body schema, sensitivity to pain etc., then he will probably find that a dose of Armstrong, like some good strong coffee, will clear his head and improve his vision. Where Armstrong may have trouble at the very outset is from those philosophers who would hesitate to use his distinction between mediate and immediate perception, and who would find his concept of sense impression insufficiently precise for his purposes. I suspect that some philosophers are also likely to say that his elucidation of, for example, pain, even if adequate as far as it goes, remains very obscure. For his elucidation contains a new concept—feeling a disturbance—which requires more explanation than he has given it. Now I cannot help thinking that some of these objections could have been avoided if he had started at another place—with the facts about, for example, pain sensitivity and reactions, and with the current views in the scientific world about what is going on here. He might then have spotted that his conclusions about pain are obviously congruent in general with the emphasis in the scientific world; and he might have picked up a few hints from this quarter about how to make his conclusions about pain less obscure. Moreover, he might also have spotted that his views can be stated, perhaps, without the aid of old world epistemological stories about immediate perception and sense impressions.

B. A. FARRELL

Henderson and Gillespie—Textbook of Psychiatry. 9th Edition Revised by Sir David Henderson & Ivor R. C. Batchelor. London: Oxford Univ. Press, 1962. Pp. 518. 42s.

No textbook which runs into nine editions can be ignored. 'Henderson & Gillespie' has held a prominent place on the reading-lists of students of psychiatry in this country since its first appearance in 1927. The present edition offers new material and some revision of the old text but the form remains recognizably the same as that of its predecessors. The book provides a readable and, on the whole, a well-balanced account of the elements of clinical psychiatry. The psychologist, however, who may be looking for something more, should be aware that although the word 'psychology' does not appear in the index the authors adhere to the tenets of the 'commonsense' psychology of Adolf Meyer, to whom the book is dedicated. In some measure the satisfaction which the critical reader will derive from the book may be related to his respect for commonsense.

MICHAEL SHEPHERD

The Fundamental Forms of Social Thought. By Werner Stark. London: Routledge & Kegan Paul, 1962. 265 pp. 32s.

Philosophical realism and nominalism as developed in two types of sociological theory, organicist and mechanist, are the 'fundamental forms of social thought' of Dr Stark's book. Organicism is traced from Aristotle and Aquinas through Spencer, Trotter, Gumplovic and Schäffle to the functionalist anthropologists. The modern model of these social thinkers is physiology. Mechanism is traced from Rousseau and Adam Smith to Menger, Bentham, Carey, Pareto, Simmel, Lundberg and Dodd. Dr Stark stresses that simple societies lend themselves to treatment as organisms, whereas the rational contractual character of economic behaviour can be analysed on the basis of physics or mechanics. There are, however, other circumstances which surely explain the different models employed in these two social sciences. The social anthropologists wanted to explain the structure of relationships, and their explanations were in terms of 'final causes' of particular social phenomena. Their data were held in a static frame because this was their experience of primitive society—they were in no position to describe, or even to recognize, the processes of change. The economists,

however, were mechanists because their concern was the processes by which an implicit logic of behaviour worked itself out in different (largely psychological) circumstances. They assumed certain basic causal relationships and wanted to articulate their operation.

Dr Stark seeks to explain the predominance of particular types of theory in particular periods. But his periods overlap and, some theorists defy classification. And less satisfactory still is the fact that different types of influence are related to these periods—sometimes the correlation is with economic conditions, sometimes with the development of particular natural sciences, and sometimes with the growth of new nations.

The book provides an interesting over-view of the development of certain aspects of social thought, though some of the theories examined are perhaps better forgotten. Other dichotomies could certainly claim to be no less fundamental—rationalist, irrationalist; evolutionist, functionalist; historicist, empiricist. It is strange to find Cooley, Tönnies and Sumner emerging as the heroes of a synthesized approach—'cultural sociology'. Cooley's dictum that the individual and society are but the distributive and collective aspects of the same thing is surely only a small stepping stone in the development of sociology.

My main disagreement with Dr Stark would concern his apparent search for 'truth' in all these theories. Is it not more useful to regard sociology as a discipline seeking explanations, and for this purpose employing models, rather than to look for the 'truth-content' in these various theories? Might we not look at organic and mechanist theories as we regard integration and coercion models of society—as heuristic constructs, to be jettisoned when they fail to help us to explain significant social phenomena?

B. R. WILSON

The Assimilation of Intellectual Refugees in Western Australia: with special reference to Hungarians. By Ronald Taft and A. G. Doczy. The Hague: Albani, 1962. Pp. 82.

This is a further study from the University of West Australia's programme on immigration, and it uses the same 'standard conceptual scheme' that appeared in the earlier studies. The limitations and weaknesses of this scheme appear even more clearly on this occasion because of the potential richness of the problem. Intellectual migrants pose the problems of cultural contact and shock much more sharply than do migrant labourers. With refugee intellectuals there is the added complication that they have now worked through their divorce from the cultural climate of their homeland and yet have to make their living within a culture that is for them alien, inferior and unchosen. Before they can start to assimilate they may have to learn to live with the fact that they cannot go back. Only then are they likely to seriously wonder whether their place of refuge is their desired place of abode.

Very little of this complexity shows through the concepts and methods of analysis used here. If one accepts the validity of the various indices, the results indicate simply that some have become satisfied, some dissatisfied with their life in Australia; some have come to identify themselves as Australians and some have come think to about things and evaluate them like Australians. From these results little seems to follow. The three indices of satisfaction, identification and acculturation bear no sensible relation to each other and are related neither to the dilemmas of the individuals about staying or going, nor to the way they have tried to resolve these dilemmas. Only the brief exercise with an alternative form of analysis based on the subjects' desires to stay or go begins to touch on these questions. However, even then the opaqueness of the indices prevents the authors and the readers from glimpsing any pattern.

The shortcomings of this report probably cannot be put down to the difficulties of the subject matter; they appear to result from using concepts that are vague and unrelated and from trying to overcome the deficiency of conceptual analysis by complicated exercises in operational definition and empirical correlation. The proliferation of items and their indiscriminate lumping together to form indices raises a host of pseudo-empirical questions, but successfully masks many real ones, and provides a much less satisfactory basis for comparative studies than do simpler straightforward indices as desire to stay, intermarriage and the like.

F. E. EMERY

Casework in Child Care. By Jean Kastell. London: Routledge & Kegan Paul, 1962. Pp. 302. 35s.

Until this Summer, 1962, there has been no book on current British casework practice in child care. This book and Noel Timms' *Casework in the Child Care Service* appeared within weeks of each other to fill a serious gap in our social work literature. It is therefore extremely disappointing that Kastell's book is, in my view, inadequate and misleading. Indeed, it is surprising that the work ever got past the publishers, for from any standpoint it is extraordinarily badly written. The style is tortuous; and there are many infelicitous phrases, vague generalizations and platitudes. The chapters are subdivided in a confused and seemingly arbitrary way and case illustrations are presented awkwardly.

Nevertheless, if the heart of the matter were present, stylistic ineptitude would be forgiven. Unfortunately, the lack of clarity in writing of casework goes far beyond mere literary incompetence and reflects a basic uncertainty about casework theory. The fundamental weakness of the book is the attempt to divide what is termed 'pure casework' from the day to day duties of the Child Care Officer. It is doubtful whether there is such a thing as 'pure casework', if by that is meant casework divorced from agency function or social setting. This author reveals particular unease about the role of the caseworker in a setting which carries authority. The impression is given that there are a wide variety of contacts with people in need of help in which the Child Care Officer is not a caseworker. In fact, although every professional social caseworker has to come to terms with the limitations and peculiar difficulties of the setting, her training as a caseworker will guide her in the way she copes with all the wide variety of situations and people which require her help. Furthermore, in implying that emotional involvement is regrettable (p. 21) there is a failure to distinguish between the over-involvement which diminishes professional competence and the deep general involvement without which no helping is possible, and in which the needs of the helper and the helped are both met.

The book deals with all major aspects of the work and gives some idea of its complexities, although the sections on 'Preventive' work and Approved School After Care are cursory. The author is clearly most at home in writing on fostering and there is much that is sound concerning this part of the work and the practice of child care. But anyone interested in the more general issues about casework must go elsewhere. Perhaps to Noel Timms' book or to the recently published book by Ferard & Hunnybun—*The Caseworker's Use of Relationships*.

OLIVE STEVENSON

Emotion: Bodily Change. K. Candal (ed.). New York and London: Van Nostrand, 1962. Pp. viii + 263. 15s.

Emotion: Bodily Change, is one of a series of paperback Insight Books to be published by van Nostrand, each devoted to a particular problem in psychology. It consists of a selection of 16 articles which originally appeared in Journals, books and the *Scientific American*. There are virtually no editorial comments but there is a useful reference list of further reading.

Papers by William James and John Watson are included as historical background. Papers by Duffy and Wenger and Jones go some way towards providing the student with an operational approach to the concept of emotion. The rest of the book is healthily devoted to experimental papers. The physiological aspects of emotion are not adequately represented; for example, the hypothalamus and limbic system are only referred to incidentally and the reticular activating system gets no mention at all.

The book is no substitute for a balanced text-book account of bodily changes in emotion, but it does provide a convenient and relatively cheap source of supplementary original readings for the student.

IAN P. HOWARD





